UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

BCI-Based Spatial Navigation Control: A Comparison Study

Mauro Rafael Oliveira Sansana

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Sinais e Imagens Médicas

Dissertação Orientada por:
Orientador: Doutor Alexandre da Rocha Freire de Andrade
Co-Orientador: Doutor Manuel João Caneira Monteiro da Fonseca

-2016-

“We are the middle children of history. Born too late to explore earth, born too early to explore space. The only last frontier at our reach is the vast ocean of our minds."

Anonymous

Acknowledgements

First of all, I would like to thank my Co-Supervisor, Manuel Fonseca, for all the patience hearing my ideas and helping me with my difficulties over these long years, and above all for always being ready to offer more help.

I would also like to thank my Supervisor, Alexandre Andrade, for providing me with some valuable knowledge when the project was at an impasse, and also for the general support throughout the project.

I must thank LaSIGE for providing me with a room with the necessary equipment, with a very special thank you to Pedro Gonçalves, and the Department of Informatics for the recording device, without which this project would not be complete.

On a more personal level, I would like to thank my parents, for understanding the need for me to extend the time I needed for this project and providing me with the means to accomplish it, and also for all the patience over these long years that I have been studying. And, of course, to my daily companion on this journey, my significant other, for her patience, dedication and for providing me with the psychological strength to never give up, no matter the difficulties.

To all my friends, thank you.

Abstract

On a general overview, current Brain-Computer Interfaces (BCIs) require significant efforts to set up, calibrate, and operate. A number of key issues have to be addressed in order to transfer BCI technology from the laboratory/clinical setting to the daily use setting.

The goal of this work is to develop a "plug and play" BCI system, using a commercially affordable portable recording device, to reduce the significant efforts of current BCIs in the fields of setup, calibration and operation. We intend to create a BCI system that allows for an easy and quick setup, with no training time ("calibrate") and no user intervention on configurations ("operate"). The secondary goal of this work involves the evaluation and comparison of two distinct systems based upon the two exogenous BCI paradigms – Sensory Evoked Potentials (SEP) and Event Related Potentials (ERP) – to identify the best paradigm for a "plug and play" BCI system.

In order to achieve these goals, we developed two distinct systems based upon the two exogenous paradigms. The SEP-based system was developed with a Steady State Visually Evoked Potential (SSVEP) paradigm, without any calibration setup, using a Power Spectral Density Analysis (PSDA) algorithm for the feature extraction system, which considers the fundamental stimuli frequencies and their first two harmonics, and a class based classification system. The ERP-based system was developed with a P300 paradigm, also without any calibration setup, using two different algorithms for the feature extraction system, the Peak Picking and Area Analysis algorithms, and the classification system used weighted voting and a class based system to combine the results from these two algorithms. These systems were then used to control a BCI-based interactive interface using only inputs from EEG signals recorded with the Emotiv® EPOC™ Headset, which allows for a very quick setup time. The interactive interface consisted of a 2D maze, in which the BCI system controls the movement in 4 directions of an avatar inside the maze.

Prior to using the developed BCI systems with the data recorded from the Emotiv headset, these were tested using freely available high quality datasets, collected with non-portable EEG recording systems in a lab environment. Afterwards, we evaluated the BCI systems using data from voluntary users collected with a commercially affordable portable recording device (i.e. Emotiv) in an out-lab environment. Unfortunately, we were unable to fully evaluate the P300-based BCI system in an Online out-lab environment in a "plug-and-play" scenario, due to the consistent low accuracy values in the Offline run.

In conclusion, considering the conditions imposed by this work namely, the out-lab environment, the commercially available recording device, and taking into account the two developed algorithms, the paradigm that obtained better performance in a “plug-and-play” scenario was the SSVEP-based BCI system, with a peak average accuracy of 89.03% and an ITR of 21.01 bits/min.

Keywords: *BCI, EEG, “plug-and-play”, SSVEP, P300.*

Resumo

De uma forma geral, as atuais Interfaces Cérebro-Computador (BCIs) requerem esforços significativos de configuração, calibração e operação. A fim de transferir a tecnologia BCI do ambiente clínico/laboratório para o ambiente de uso diário, ainda é necessário abordar um certo número de questões fundamentais. Estas questões incluem a facilidade e conveniência da utilização diária, o aspeto estético, a segurança, a fiabilidade, a utilidade das aplicações BCI na vida diária do utilizador, e a necessidade de supervisão técnica especializada continuada. O custo de suporte técnico contínuo pode ser alto, e esse apoio por vezes só pode estar disponível a partir de alguns grupos de pesquisa. Outro possível obstáculo para mover a tecnologia BCI para um uso mais prático é a demanda de atenção do utilizador, pois tem sido reportada uma rápida fadiga de utilizadores em alguns estudos de controlos de BCIs. Algumas aplicações BCIs atuais requerem uma série exaustiva de comandos, o que contribui ainda mais para o aumento da fadiga do utilizador. Portanto, o desenvolvimento de sistemas BCI com complexidade reduzida, tempo de resposta rápido, fácil configuração e necessidade mínima de suporte técnico é essencial para a ampla disseminação da tecnologia BCI.

O objetivo do presente trabalho é desenvolver um sistema BCI "plug-and-play", usando um dispositivo portátil de gravação comercialmente disponível, de forma a reduzir os esforços significativos das BCIs atuais nas áreas da configuração, calibração e operação. Noutras palavras, pretende-se criar um sistema BCI que permita uma configuração fácil e rápida, sem tempo de treino ("calibração") e sem intervenção do utilizador em configurações ("operação"). O objetivo secundário deste trabalho envolve a avaliação e comparação de dois sistemas distintos, construídos com base nos dois paradigmas exógenos da tecnologia BCI – Potenciais Sensorialmente Evocados (SEP) e Potencias Relacionados com Eventos (ERP) – de forma a identificar o melhor paradigma para um sistema BCI "plug-and-play". Para podermos comparar estes dois sistemas, e testar o sistema como um todo, desenvolvemos ainda uma interface interativa para navegação espacial, ou seja, um labirinto, e estudámos a Taxa de Transferência de Informação (ITR) e a fadiga relativa dos diferentes utilizadores.

Para atingir estas metas, desenvolvemos dois sistemas distintos com base nos dois paradigmas exógenos. O sistema baseado em SEP foi desenvolvido com um paradigma de *Steady State Visually Evoked Potential* (SSVEP), sem qualquer configuração de calibração. O sistema BCI baseado em SSVEP foi construído utilizando um algoritmo de Análise de Potência da Densidade Espectral (PSDA) para o sistema de extração de caraterísticas, que considera as frequências fundamentais de estimulação e as suas primeiras duas harmónicas, e um sistema de classificação baseado em classes. O sinal de EEG foi recolhido a partir dos elétrodos $O1$ e $O2$. O sistema baseado em ERP foi desenvolvido com um paradigma P300, que geralmente requer alguma configuração de calibração, mas para os propósitos deste estudo, também foi desenvolvido sem qualquer configuração de calibração. O sistema baseado em P300 foi construído utilizando dois algoritmos diferentes para o sistema de extração de caraterísticas, os algoritmos de Escolha de Picos e de Análise de Áreas, e o sistema de classificação utilizou um sistema de votação com pesos e um sistema baseado em classes de forma a combinar os resultados destes dois algoritmos. Os sinais EEG recolhidos dos elétrodos $F3, FC5, FC6$ e $F4$ foram utilizados para controlo da interface, e os sinais dos restantes elétrodos foram utilizados para a construção de um filtro espacial de Referência Média Comum (CAR). Estes sistemas BCI foram depois utilizados para controlar uma interface interativa baseada em BCI utilizando apenas os sinais de EEG recolhidos com o *Headset* Emotiv® EPOC™ como *input*, o que permite um tempo de preparação rápido. A interface interativa consistiu num labirinto 2D, no qual o sistema BCI controla o movimento em quatro direções de um avatar dentro do labirinto. Este controlo é conseguido através da existência de quatro setas intermitentes distintas umas das outras, e para cada paradigma. Para os utilizadores escolherem a direção que pretendem que o avatar se mova, simplesmente têm de olhar para a seta que representa essa direção, e o sistema irá reconhecer a intenção do utilizador, minimizando as dificuldades de operação.

Antes de utilizar os sistemas BCI desenvolvidos para analisar os dados recolhidos com o *Headset* Emotiv, estes foram testados utilizando um conjunto de dados ("*dataset"*) de alta qualidade livremente disponíveis, e recolhidos com sistemas de gravação de EEG não portáteis num ambiente de laboratório. O sistema de BCI P300 foi o algoritmo que obteve a maior precisão média (87.09%), seguido do sistema BCI SSVEP com uma precisão média máxima de 77.25%.

Avaliámos ainda os mesmos algoritmos, utilizando os dados de dez utilizadores voluntários saudáveis, quatro femininos e seis masculinos idade média de 23.7 ± 3.65 anos, recolhidos com um dispositivo de gravação portátil comercialmente acessível (ou seja, o Headset Emotiv) num ambiente fora do laboratório. A fase Online SSVEP foi a que obteve a precisão média mais alta (89.03%), seguida da fase Offline SSVEP com uma precisão semelhante (87.50%) e finalmente a fase Offline P300 com a precisão média menor (28.13%). A ITR alcançada pelo sistema BCI SSVEP Online foi de 21.01 bits/min.

Finalmente, e através da análise dos questionários de autoavaliação preenchidos pelos utilizadores durante o estudo, observámos que o nível de concentração do utilizador tem uma relação direta ($\rho = 0.8725; r^2 = 0.7612; p = 0.0010$) com a precisão do sistema BCI SSVEP Online, isto é, quanto menor a precisão do sistema, mais intensamente o utilizador tem de se concentrar para terminar o labirinto e vice-versa.

Infelizmente, não nos foi possível realizar uma avaliação completa do sistema BCI baseado em P300 num ambiente Online fora do laboratório com dados de utilizadores voluntários, recolhidos a partir de um dispositivo de gravação portátil comercialmente acessível (cenário "plug-and-play"), devido aos baixos valores de precisão observados consistentemente na corrida Offline (média de 28.13%). Apesar disto, acreditamos que o sistema BCI P300 que desenvolvemos é suficientemente robusto para ser capaz de extrair e classificar sinais EEG que contenham eventos P300, isso pode ser verificado pelo valor de precisão alcançado pelo sistema durante a avaliação usando os conjuntos de dados de P300, o que nos leva a acreditar que o problema pode estar no sistema de entrega de estímulos ou no dispositivo de gravação, devido ao facto de este dispositivo não ter na sua configuração as posições de elétrodos recomendadas para a recolha de sinais de EEG que contenham eventos P300.

Em conclusão, considerando as condições impostas por este trabalho, nomeadamente, o ambiente fora de laboratório, o dispositivo de gravação portátil comercialmente acessível, e tendo em conta os dois algoritmos desenvolvidos, o paradigma que obteve a melhor performance num cenário "plug-and-play" foi o sistema BCI baseado em SSVEP, com uma precisão média máxima de 89.03% e uma ITR de 21.01 bits/min.

Palavras-chave: *BCI, EEG, "plug-and-play", SSVEP, P300.*

Table of Contents

List of figures

List of Tables

Abbreviations

AEP	Auditory Evoked Potentials
BCI	Brain-Computer Interface
BMI	Brain-Machine Interface
CAR	Common Average Referencing
CCA	Canonical Correlation Analysis
DFT	Discrete Fourier Transform
DNI	Direct Neural Interface
ECG	Electrocardiography
ECoG	Electrocorticography
EEG	Electroencephalography
EMG	Electromiography
EOG	Electrooculography
ERD	Event Related Desynchronization
ERP	Event Related Potentials
ERS	Event Related Synchronization
FFT	Fast Fourier Transform
fMRI	Functional Magnetic Resonance Imaging
ISI	Inter Stimulus Interval
ITI	Inter-Trial Intervals
ITR	Information Transfer Rate
LDA	Linear Discriminant Analysis
LFP	Local Field Potentials
MEC	Minimum Energy Combination
MEG	Magnetoencephalography
MI	Motor Imagery
MMI	Mind-Machine Interface
NIS	Near Infrared Spectroscopy
PSD	Power Spectral Density
PSDA	Power Spectral Density Analysis

PTB-3	Psychophysics Toolbox Version 3
RGB	Red Green Blue
RGBA	Red Green Blue Alpha
SCP	Slow Cortical Potentials
SEP	Sensory Evoked Potentials
SMR	Sensorimotor Rhythms
SNR	Signal-To-Noise Ratio
SSEP	Steady State Evoked Potential
SsEP	Somatosensory Evoked Potential
SSVEP	Steady State Visually Evoked Potential
STR	Single Trial Recognition
SVM	Support Vector Machines
SWLDA	Step-Wise Linear Discriminant Algorithm
UI	User Interface
VEP	Visual Evoked Potentials
VGA	Video Graphics Array
VR	Virtual Reality

1. Introduction

In this chapter we present the motivations and problems related to this work, our main goals and the developed solution to address the presented problems. We also enumerate the main contributions and results of our work and present an outline of this document.

1.1. Background, Motivations and Problem

A Brain-Computer Interface (BCI) is a direct pathway of communication between the brain and an external device. The need for the creation of BCI-based devices arose from the wish to create alternative options for control and communication, with a special focus on individuals with severe disabilities. However, more recently, their potential was extended to the domain of treatment for neurological diseases, monitoring of the brain status and videogames [1].

The most practical and largely used BCIs are those based on non-invasive recording of signals from Electroencephalography (EEG) directly on the scalp [2].

BCIs are driven by a variety of neurophysiologic signals that reflect in vivo brain activity. Depending on the biophysical nature of the signal source, these signals can be broadly grouped into three categories: electrophysiological, magnetic, and metabolic [3]. The use of magnetic and metabolic signals for BCI systems in a real-life daily use is currently not realistic, due to the high cost and size of the equipment and technical difficulties involved [4]. Therefore, the main source of signals for BCI systems are electrophysiological ones.

Brain-computer interaction is usually slower and less accurate than most modalities that are currently available for computer interaction, e.g. keyboards, mice, gamepads and joysticks. Furthermore, BCIs often require a lot of training and setup time. The information transfer rate (ITR) of BCIs is still around up to 25 bits per minute [1], which is incomparable with keyboard speeds of over 300 characters per minute, roughly 2400 bits per minute.

EEG-based BCIs are currently adaptable for practical independent use by disabled people outside of the laboratory [5] and are in fact currently in use by a small number of paralyzed people at home in their daily lives [6]. Despite this achievement, continuing development of practical EEG-based BCI systems is needed to address existing issues.

On a general overview, all current BCIs require significant efforts to set up, calibrate, and operate [7]. The kind and degree of effort vary substantially across BCIs. Local Field Potentials (LFP)-, spike-, or Electrocorticogram (ECoG) -driven BCIs require surgery and constant monitoring of the scalp-port through which the wires run to the electrodes. EEG electrodes, which can be applied in a few minutes, require periodic reapplications. Daily recalibration may be needed (particularly for intracortical BCIs). A number of key issues have to be addressed in order to transfer BCI technology from the laboratory/clinical setting to the daily use setting [8]. These include the ease and convenience of daily use, cosmesis, safety, reliability, usefulness of the BCI applications in the user's daily life, and the need for ongoing expert technical oversight. The cost of ongoing technical support may be high, and such support may only be available from a few research groups.

Another possible obstacle to moving BCI technology into practical use is their demand on the user's attention. Rapid fatigue of users has been previously reported in certain studies of BCI control [9], [10], and inconsistent performance by individual users is characteristic of most methods. Ongoing changes in the user's performance (due to fatigue, distraction, disease progression, etc.) require continuing adaptation of the BCI system. Moreover, some BCI applications currently require an exhaustive series of user commands (e.g., different mental activation tasks), which further contribute to an increase in user fatigue.

Therefore, the development of BCI systems with reduced complexity, fast response time, easy setup and minimal need for ongoing technical support is essential for the widespread dissemination of BCI technology.

1.2. Goals

The goal of this work is to develop a "plug and play"[1] BCI system, using a commercially affordable portable recording device, to reduce the significant efforts of current BCIs in the fields of setup, calibration and operation. In other words, we intend to create a BCI system that allows for an easy and quick setup, with no training time ("calibrate") and no user intervention on configurations ("operate").

The secondary goal of this work involves the evaluation and comparison of two distinct systems based upon the two exogenous BCI paradigms – Sensory Evoked Potentials (SEP) and Event Related Potentials (ERP) – to identify the best paradigm for a "plug and play" BCI system. In order to compare these two paradigms, and test the whole system, we need to develop an interactive interface for spatial navigation, i.e. a maze, and study the ITR and relative fatigue of different test subjects.

[1] The term "plug and play" has several meanings depending on the context of usage. In the context of this work, it is used to describe a system in which the user has minimal interaction on the setup of the system itself, and therefore the only interaction that is intended is for the user to "plug" the recording device and "play" with the interactive interface.

1.3. Developed Solution

To achieve these goals, we developed two distinct systems based upon the two exogenous paradigms.

The SEP-based system was developed with a Steady State Visually Evoked Potential (SSVEP) paradigm, without any calibration setup. The SSVEP-based BCI system was built using a Power Spectral Density Analysis (PSDA) algorithm for the feature extraction system, which considers the fundamental stimuli frequencies and their first two harmonics, and a class based classification system. The EEG signal was collected from electrodes $O1$ and $O2$.

The ERP-based system was developed with a P300 paradigm, which usually requires some calibration setup, but for the purposes of this study, was developed without any calibration setup. The P300-based BCI system was built using two different algorithms for the feature extraction system, the Peak Picking and Area Analysis algorithms, and the classification system used weighted voting and a class based system to combine the results from these two algorithms. The EEG signal collected from electrodes $F3, FC5, FC6$ and $F4$ was used for control, and the signal from the remaining electrodes was used to construct a Common Average Referencing (CAR) filter.

These systems were then used to control a BCI-based interactive interface using only inputs from EEG signals recorded with the Emotiv® EPOC™ Headset, which allows for a quick setup time. The interactive interface consisted of a 2D maze, in which the BCI system controls the movement in 4 directions of an avatar inside the maze. This control is achieved by having 4 arrows flashing in a distinct way from each other, and for each paradigm. For users to choose the direction they want the avatar to move to, they simply have to gaze at the arrow that represents that direction, and the system will recognize the users' intention, minimizing the operation difficulties.

1.4. Contributions

This work presents a "plug-and-play" solution based on two innovative SSVEP and P300 algorithms, without any training time. We also present a comparative study of both algorithms, with high quality datasets collected from an in-lab environment using analog EEG systems, and also with data collected from voluntary users in an out-lab environment using a portable recording device.

Finally, in order to perform this last comparative study, we also developed an easily expandable and integrated interactive interface, that works together with the portable recording device to offer a fully BCI controlled user interface inside a MATLAB environment.

1.5. Outline

In the second chapter we introduce the basic concepts necessary to understand this work, the main components and paradigms of a BCI system and we also discuss the common evaluation practices in BCIs. In the third chapter we analyse and discuss the most relevant and recent work that has been developed over the years on the field of BCIs, comparing them with our proposed solution. In the fourth and fifth chapter, we present with detail the solution we used to achieve our goals. In the sixth chapter we present and discuss the experimental plan used to evaluate the proposed solution, and we analyse the results obtained. Finally, in the seventh chapter we summarize the main contributions of our work and discuss future work.

2. Background Theory

A brain–computer interface (BCI), sometimes called a mind-machine interface (MMI), direct neural interface (DNI), or brain–machine interface (BMI), is a direct communication pathway between the brain and an external device.

BCIs are often directed at assisting, augmenting, or repairing human cognitive or sensory-motor functions.

Research on BCIs began in the 1970s at the University of California, Los Angeles (UCLA) under a grant from the National Science Foundation (NSF), followed by a contract from the Defense Advanced Research Project Agency or DARPA [11], [12]. The papers published after this research also mark the first appearance of the expression Brain–Computer Interface in scientific literature.

In the following subsections, we introduce the concept of Electroencephalography and its importance in BCI systems, present the main modules associated with BCI systems as well as the main paradigms and discuss evaluation methods for BCI systems.

2.1. Electroencephalography

Electrophysiological signals can be classified according to the degree of invasiveness of the recording device [3], [8]:

- **Non-invasive systems** record signals without directly interfering with the skull integrity (e.g. Electroencephalography or EEG). Since EEG is non-invasive (scalp measure), it provides the simplest and safest BCI recording method;
- **Cortical surface systems** record signals using electrodes surgically placed on the surface of the cortex (e.g. Electrocorticography or ECoG). ECoG provides essentially the same signals as EEG, but because they are recorded in close proximity to the brain, they are not affected by the insulation features of the skull and dura mater;
- Finally, **intracortical systems** record electrical activity from electrodes implanted deeply in the brain.

Both ECoG and Intracortical recordings provide wider frequency range, higher topographical resolution, and better signal quality and dimensionality than EEG. However, both methods require surgery, which brings both ethical and health issues [8], [13].

Researchers tend to prefer non-invasive systems due to their safety and easier implementation. It is also important to note that since BCI is working towards an approach to the common user, non-invasive systems are a perfectly logic choice and therefore they should become the main focus of BCI systems to the "large masses" research throughout the next years. The methods available on this class range from Functional Magnetic Resonance Imaging (fMRI) to Near Infrared Spectroscopy (NIS) and Magnetoencephalography (MEG) signals. However the most widely used non-invasive BCI systems are by far those based on EEG signals, where acquisition of the signal is done using electrodes placed on the scalp [8], [14]–[16].

EEG can then be further defined as a non-invasive method to record electrical activity of the brain along the scalp. EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain [17].

The measurement of EEG signals can be made in many ways, ranging from dry electrodes placed directly on the scalp – that collect the signal to be then filtered, sampled and digitized with analogical devices – to more advanced portable commercial systems – that record, filter, sample and digitize the signal directly on the device itself.

High fidelity signals from bioelectrodes are of paramount importance for medical applications such as EEG, Electromyography (EMG) and Electrocardiography (ECG). The most commonly used bioelectrode is the silver/silver chloride (Ag/AgCl) type which can be found in reusable form, but is most often used as a disposable electrode [18]. This type of electrodes is often referred to as a wet electrode type, due to its dependency on an electrolytic gel to form a conductive path between the skin and the electrode. Ag/AgCl electrodes have disadvantages in some applications which has led to the investigation of alternative electrode designs, such as the dry electrodes and insulating electrodes. Dry electrodes consist of a benign metal (e.g. stainless steel) or polymer used without the need of an electrolyte gel between electrode and skin. Insulating electrodes consist of a metal or semiconductor with a thin dielectric surface layer, so that the bioelectric signal is capacitively coupled from the skin to the substrate.

Reasons given for past investigation of alternatives to gel-based electrodes for bioelectric applications include:

- **The performance regarding power lines interference of wet types is limited by electrode/skin contact impedance.** The use of active electrodes (where buffering/amplification takes place at the electrode site) provides much less emphasis on the skin–electrode impedance [19];
- **The use of an electrolyte is inconvenient.** For chronic use, the reliance on an electrolyte leads to reduced signal quality as the gel [20] and the reapplication of gel may not be feasible. The recording may be in a sensitive area, e.g. near the eyes [21], previous skin treatment may disable standard electrodes, such as the use of petroleum based burns ointment [22], or the spacing required between electrodes may be so small that smearing of the electrolyte, and thus short circuiting of the bioelectric signal, may occur [23], [24]. Additionally, and perhaps most importantly, the application and removal of electrolyte gels is an unpleasant process for the subject, and time consuming for the clinician or carer [20], [23], [25].
- **There are toxicological concerns with electrolyte gels.** Although dermatitis from gels used in bioelectric recordings is 'exceedingly rare' [26], concerns regarding dermatological responses are common [20], [22], [24].

In conventional scalp EEG, the recording is obtained by placing electrodes on the scalp with a conductive gel or paste (i.e. wet electrodes), usually after preparing the scalp area by light abrasion to reduce impedance due to dead skin cells. Many systems typically use electrodes, each of which is attached to an individual wire. Some systems use caps or nets into which electrodes are embedded; this is particularly common when high-density arrays of electrodes are needed. Each electrode is connected to one input of a differential amplifier (one amplifier per pair of electrodes); a common system reference electrode is connected to the other input of each differential amplifier. These amplifiers amplify the voltage between the active electrode and the reference (typically $1{,}000 - 100{,}000$ times, or $60 - 100$dB of voltage gain). In analog EEG, the signal is then filtered and the EEG signal is output as the deflection of pens as paper passes underneath. Most EEG systems these days, however, are digital, and the amplified signal is digitized via an analog-to-digital converter, after being passed through an anti-aliasing filter. Analog-to-digital sampling typically occurs at 256 to 512Hz in clinical scalp EEG; and sampling rates of up to 20kHz are used in some research applications. The digital EEG signal is stored electronically and can be filtered for display. Typical settings for the high-pass filter and a low-pass filter are $0.5 - 1$Hz and $35 - 70$Hz, respectively. The high-pass filter typically filters out slow artifacts, such as electrogalvanic signals and movement artifacts, whereas the low-pass filter filters out high-frequency artifacts, such as EMG signals. An additional notch filter is typically used to remove artifacts caused by electrical power lines (60Hz in the United States and 50Hz in many other countries) [17]. This kind of system may be adequate in a research scenario, but it is not practical for daily usage.

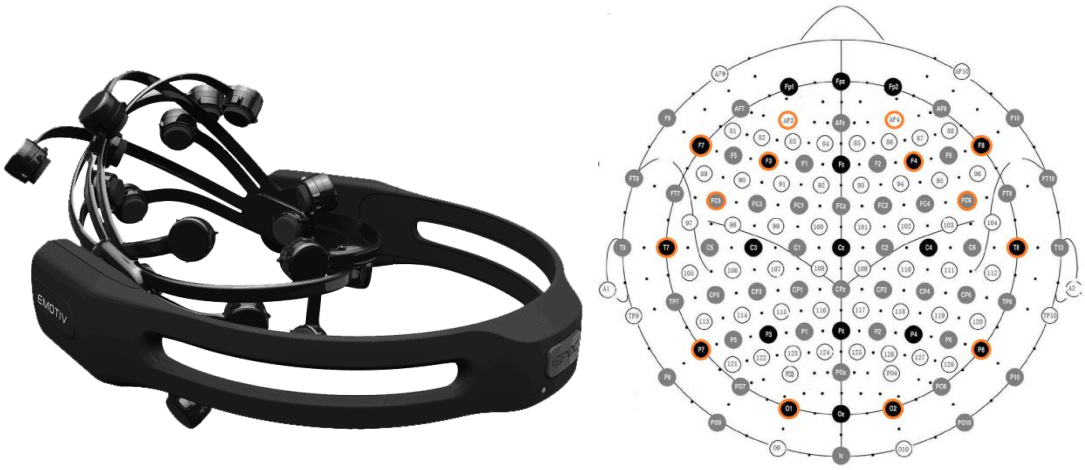


Figure 2.1 - The Emotiv® EPOC+™ Headset (on the left[2]) and the international standard electrode montages (on the right[3]).The image on the right represents the 10-20 (Black circles), 10-10 (10-20 + Gray circles) and 10-5 (Black dots and Open circles) electrode montages. The electrodes used by the headset are highlighted in orange and the names of the channels are (from top to bottom and left to right): AF3, AF4, F7, F3, F4, F8, FC5, FC6, T7, T8, P7, P8, O1, and O2.

In order to make BCI-based system more portable and easy to use, some companies have developed portable commercial systems that record, filter, sample and digitize the signal directly on the device itself, thus removing the need to have big setup apparatus and brings the BCI technology closer to the everyday user.

One example of these portable commercial systems is the Emotiv® EPOC+™ Headset, shown in Figure 2.1. This headset is a device that uses 16 saline soaked felt pads electrodes. These sensors use dry electrode technology through the usage of gold plated contacts, but also make use of a saline solution in place of the common electrode gel found on the wet electrode technology. The 16 electrodes are divided into 14 sensors plus 2 references. The Emotiv has a 2048 internal sample rate, that is filtered and downsampled, to 128 samples per second per channel, with a 14 bit resolution per channel and has a wireless 2.4GHz USB-based receiver for data connectivity. Furthermore, this device also as a built-in software for detection of various facial expressions, emotional states and some mental commands [27]. Compared with other wireless BCI devices (e.g. g.tec), the Emotiv® EPOC+™ Headset has a lower price and shorter preparation time prior to use [28].

In order to ensure standardization and reproducibility amongst studies using EEG signals, it was created the $10-20$ system for scalp electrodes location. The "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front–back or right–left distance of the skull.

[2] Adapted from https://emotiv.com/store/epoc-detail/
[3] Adapted from http://emotiv.wikia.com/wiki/Emotiv_EPOC

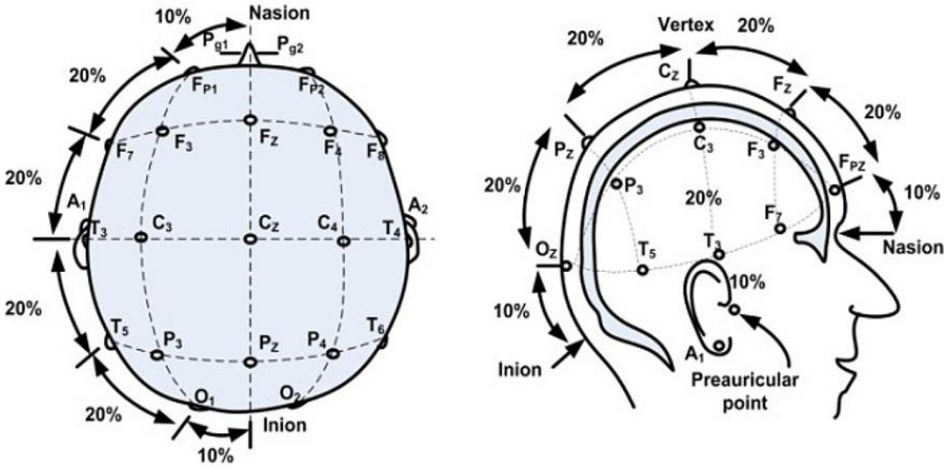


Figure 2.2 – Illustrative scheme depicting the 10-20 system for electrode placement on the scalp. The image on the left is a top view and the image on the right is a lateral view, from the right. Adapted from [29].

This system describes the location of scalp electrodes in the context of an EEG test, and is based on the relationship between the location of an electrode and the underlying area of the cerebral cortex. Each site has a letter to identify the lobe and a number to identify the hemisphere location. The letters F, T, C, P and O stand for frontal, temporal, central, parietal, and occipital lobes, respectively. Note that there exists no central lobe; the "C" letter is used only for identification purposes. A "z" (zero) refers to an electrode placed on the midline. Even numbers $(2, 4, 6, 8)$ refer to electrode positions on the right hemisphere, whereas odd numbers $(1, 3, 5, 7)$ refer to those on the left hemisphere. In addition, the letter codes A, Pg and Fp identify the earlobes, nasopharyngeal and frontal polar sites respectively [30]. In Figure 2.2, we can see the various electrodes, the spacing between them and their global position on the scalp.

2.2. BCIs – Common Modules & Paradigms

On a more technical level, a BCI can be perceived as a combination of various elements that work together in order to translate brain signals into commands. In general, the design of a BCI consists of a combination of these three essential elements [1]:

- **Signal acquisition & Pre-processing**. As discussed earlier, the most common method of signal acquisition for BCI consists on EEG signals. Usually, this element also implies that the design of the BCI must take into account the channels to use, according to the intended application of the system. This element may also include some pre-processing of the signal from the selected channels (i.e. filtering and noise reduction);

- **Feature extraction & Classification**. This element extracts signal features that should have strong correlations with the user's intent, that is, they encode the intent of the user. The signal features extracted can be in the time or frequency domain, or both. The most common signal features used in current BCI systems include amplitudes or latencies of event-evoked potentials (e.g. P300), frequency power spectra (e.g. sensorimotor rhythms), or even firing rates of individual cortical neurons [3]. These features will, in turn, be classified accordingly to the user's intent;
- **Translation/Device output**. This element converts the extracted signal features into device commands via translation algorithms. The signal feature thus extracted and translated provides the output to operate an external device, communicate with an interface or control a virtual environment.

Depending upon the control tasks, EEG-based BCIs can be divided into 4 distinct paradigms: "Sensory Evoked Potentials", "Event Related Potentials", "Event Related De/Synchronization" and "Slow Cortical Potentials". The first two paradigms further belong to the category of exogenous paradigms, in which the user focuses attention on a set of stimuli, which produce an autonomic response that is detected by the BCI. The last two belong to the category of endogenous paradigms, in which the user voluntary performs a mental task, such as imagined movement, to create changes in brain signals that can be detected by the BCI. Endogenous responses do not require a stimulus, although prompts and cues may be used to improve response characteristics [31]. In the next subsections we describe the various paradigms associated with EEG-based BCI systems.

2.2.1. Sensory Evoked Potentials

An evoked potential is an electrical potential signal recorded from the nervous system following the presentation of a stimulus. Sensory evoked potentials (SEP) are recorded from the central nervous system following the stimulation of a sense organ. Evoked potential amplitude tends to be low, and often requires EEG signal averaging techniques to be differentiated from other biological signals and ambient noise [32]. There are three kinds of evoked potentials in widespread clinical use: i) auditory evoked potentials (AEP), which are elicited by a click or tone stimulus presented through earphones, they are usually recorded from the scalp but originate at brainstem level; ii) visual evoked potentials (VEP), which are elicited by a flashing light or changing pattern on a monitor; iii) somatosensory evoked potentials (SsEP), which are elicited by electrical stimulation of peripheral nerve [33].

In the case of the elicitation of an evoked potential, the brain attains a steady-state regime in which the amplitude and phase of the frequency components (i.e. harmonics) of the response are approximately constant over a long period of time [34]. This 'constant' characteristic is not a time-domain observation. What remains constant is the spectral distribution (in the frequency domain), not the raw EEG amplitude in time domain. In this case, the signal is called a Steady State Evoked Potential (SSEP) which contains stationary periodic oscillations. The SSEP signal itself slightly fluctuates in amplitude in the time domain, amplitude even tends to decrease after the first few oscillations.

Consequently, the amplitude distribution of the spectral content of SSEP, with characteristic SSEP peaks, remains stable over time. Because these characteristics are constant, many applications can be derived from SSEP propagation properties [35].

The visual system has received much attention in neuroscience during the last century. From retinal photoreceptors, visual perceptions propagate first to visual areas, and next to the rest of the brain (Figure 2.3) [36]. Following the presentation of visual stimuli, **Steady State Visually Evoked Potentials** (SSVEPs) can be recorded in the visual areas, more precisely near the striate cortex. SSVEPs are signals that are natural responses to visual stimulation at specific frequencies.

SSVEPs are useful in research because of the excellent signal-to-noise ratio [37] and relative immunity to artifacts [32]. Due to SSVEPs particular spatial distribution, the electrodes from the 10-20 system normally used to record this type of signal are those from the Occipital (O) and ParaOccipital (PO) areas.

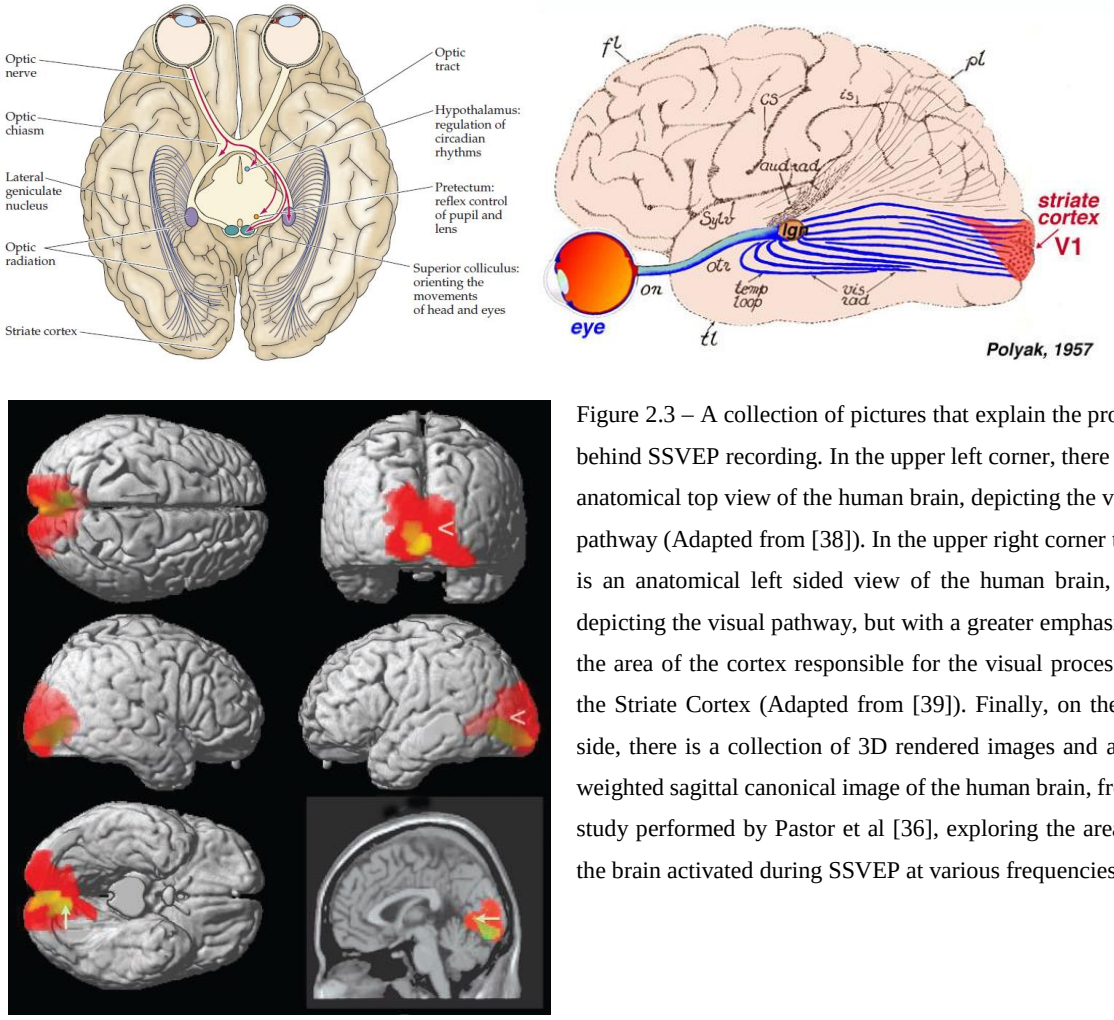


Figure 2.3 – A collection of pictures that explain the process behind SSVEP recording. In the upper left corner, there is an anatomical top view of the human brain, depicting the visual pathway (Adapted from [38]). In the upper right corner there is an anatomical left sided view of the human brain, also depicting the visual pathway, but with a greater emphasis on the area of the cortex responsible for the visual processing, the Striate Cortex (Adapted from [39]). Finally, on the left side, there is a collection of 3D rendered images and a T1-weighted sagittal canonical image of the human brain, from a study performed by Pastor et al [36], exploring the areas of the brain activated during SSVEP at various frequencies.

When the retina is excited by a visual stimulus ranging from 3.5 Hz to 75 Hz, the brain generates electrical activity at the same (or multiples of the) frequency of the visual stimulus [40]. The most common visual stimuli used are represented below in Figure 2.4.

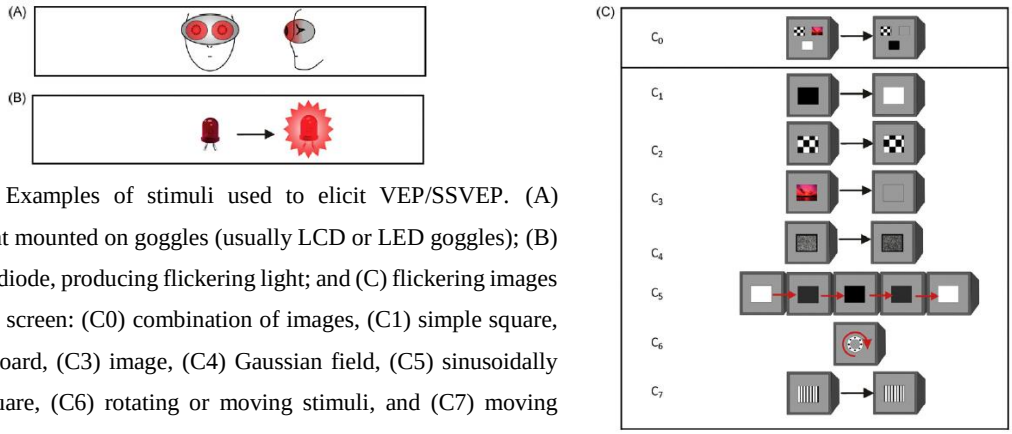


Figure 2.4 – Examples of stimuli used to elicit VEP/SSVEP. (A) Flickering light mounted on goggles (usually LCD or LED goggles); (B) light-emitting diode, producing flickering light; and (C) flickering images on a computer screen: (C0) combination of images, (C1) simple square, (C2) checkerboard, (C3) image, (C4) Gaussian field, (C5) sinusoidally modulated square, (C6) rotating or moving stimuli, and (C7) moving vertical or horizontal gratings. Adapted from [35].

2.2.2. Event Related Potentials

Event related potentials (ERPs) are the measurement of the brain responses to specific cognitive, sensory or motor events [41]. ERP waveforms consist of a series of positive and negative voltage deflections, which are related to a set of underlying components [42]. Though some ERP components are referred to with acronyms, most components are referred by a letter (N-negative; P-positive) indicating polarity, followed by a number indicating either the latency in milliseconds or the component's ordinal position in the waveform (Figure 2.5).

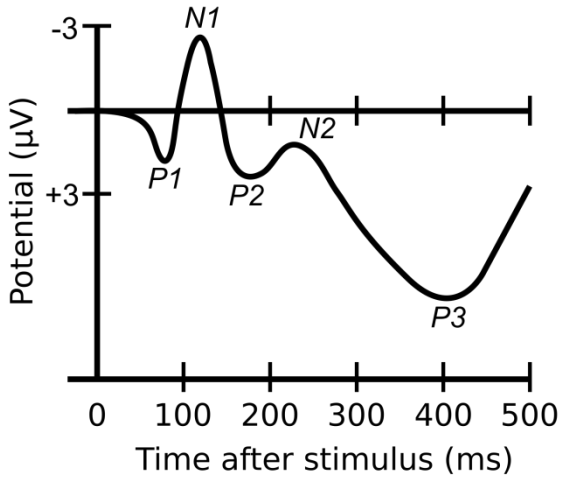


Figure 2.5 – A graphical representation of an EEG waveform showing several ERP components. Note that this waveform is plotted with the negative voltages upward, a common practice in ERP research. Adapted from [42].

P300 is a major peak of an ERP waveform, and one of the most used components of an ERP [41]. The P300 ERP often is elicited with a two stimulus discrimination task. This procedure has been dubbed the 'oddball paradigm', since two stimuli are presented in a random series such that one of them occurs relatively infrequently, i.e. the oddball. The presentation of a stimulus in an oddball paradigm can produce a positive peak in the EEG, $300msec$ after onset of the stimulus, which can be either visual, auditory or somatosensory [43].

Although the oddball task has been used most often in basic ERP studies, the P300 potentials can also be obtained from a 'single-stimulus' paradigm, by replacing the standard stimuli (non-oddball) with 'silence' [44].

The spatial amplitude distribution of P300 is strongest in the occipital region of the brain and is symmetric around central location Cz (10-20 system) [41]. The electrodes typically used to record P300 are shown in Figure 2.6.

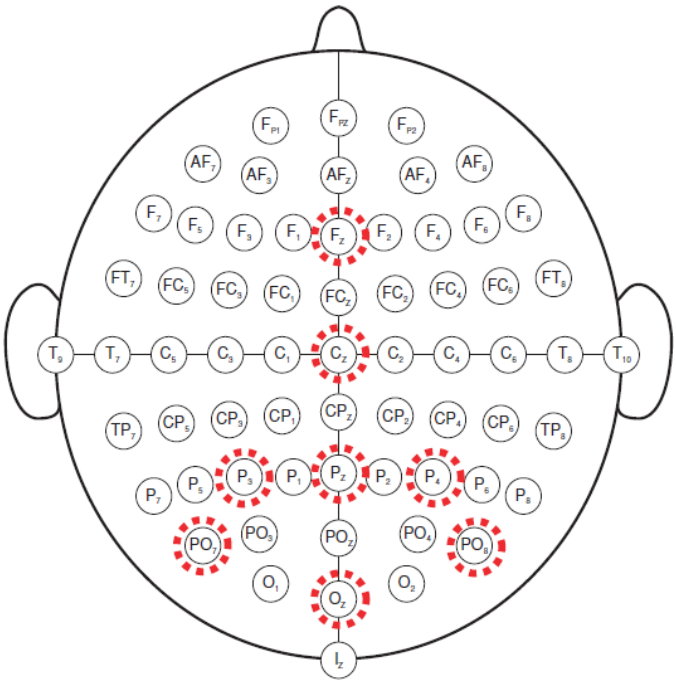


Figure 2.6 – Electrode montage typically used in P300-based BCIs. The eight electrodes used are indicated by the dotted red circles. Adapted from [45].

2.2.3. Event Related De/Synchronization

The traditional frequency range of EEG signal varies from $4Hz$ to $100Hz$ depending on the brain activity and physiopsychological state of the subject, with an ever-increasing interest in very high frequencies beyond 100 Hz. There are six different bands in this frequency range (Table 2.1) [17].

Table 2.1 – EEG rhythms, their frequency bands and common location. Adapted from [17].

Rhythm	**Frequency Band**	**Location**
Delta (δ)	$1-4Hz$	Frontally in adults
Theta (θ)	$4-7Hz$	Found in locations not related to task at hand
Alpha (α)	$8-13Hz$	Central sites ($C3-C4$) at rest
Mu (μ)	$8-13Hz$	Sensorimotor cortex
Beta (β)	$13-30Hz$	Both sides, more evident frontally
Gamma (γ)	$>30Hz$	Somatosensory cortex

In general, the frequency of brain oscillations is negatively correlated with their amplitude, which means that the amplitude of fluctuations decreases with increasing frequency. For example, the μ rhythm with a frequency between 8 and $13Hz$ has a larger amplitude than the central beta β rhythm with frequencies around $20Hz$. The β rhythm again has a larger amplitude than oscillations around $40Hz$ [46].

μ rhythm, overlapping with the α rhythm, is generated by motor and sensorimotor cortex, is eminent when subjects are resting and is reduced by self-initiated movement and observed movement. Electrodes placed overlying sensorimotor cortex can record EEG signals from the range of $8Hz$ to $30Hz$ (α, β and μ waves). By analysing this range of EEG signal, it is possible to find a one to one correspondence between the EEG signals and various motor imagery (MI) tasks performed by a human [47].

Changes in μ and β rhythm amplitudes are referred to as event-related desynchronization (ERD) (i.e. decrease) and event-related synchronization (ERS) (i.e. increase). Typically, changes in μ and β rhythms are associated with movement, sensation, and MI (i.e. imagination of movement). The rhythms decrease or desynchronize with movement or its preparation, and increase or synchronize after movement and with relaxation [17].

Sensorimotor rhythms (SMRs) are recorded over sensorimotor cortex (Figure 2.7), typically with the electrodes from the Central area ($C3$ and $C4$).

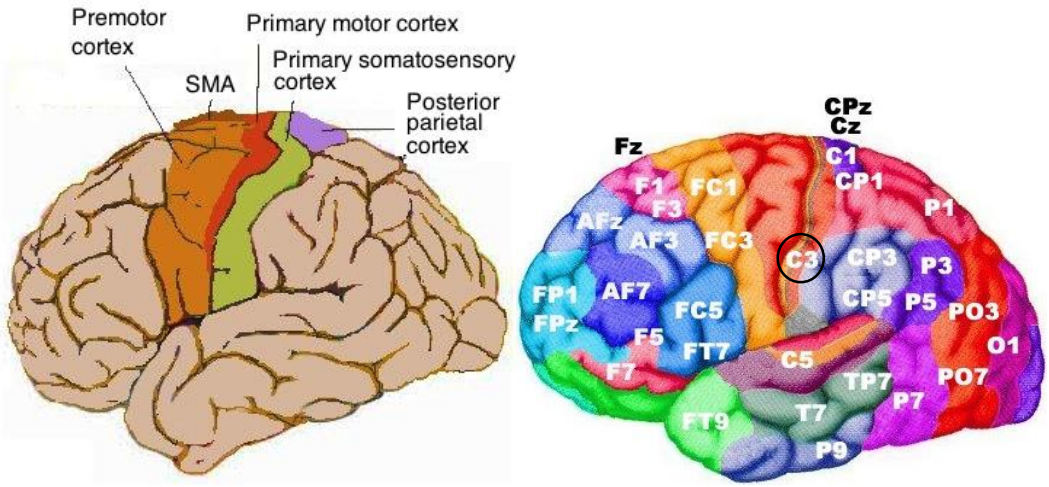


Figure 2.7 – Topography of the Human motor cortex (on the left[4]) and corresponding $10-10$ system electrodes (on the right[5]). Note that $C3$ (black circle) is located in the primary somatosensory cortex.

The meaning of SMRs is still not fully understood. Phenomenologically, a person is producing stronger SMR amplitude when the corresponding sensory-motor areas are idle, e.g. during states of immobility, and SMRs typically decrease in amplitude when the corresponding sensory or motor areas are activated, e.g. during motor tasks and even during motor imagery [17].

2.2.4. Slow Cortical Potentials

Slow Cortical Potentials (SCPs) are slow voltage changes in the cortex. They occur over $0.5-10.0$s and are among the lowest frequency features of EEG. Negative shifts of SCPs represent cortical activation associated with movement or other functions, while positive SCP shifts accompany reduced cortical activation [3], [48], [49]. With extensive training, sometimes months, the user learns to control SCP positive or negative voltage shifts. Due to the extensive period of training and very slow rate of communication, this type of signal is not commonly used in BCIs.

[4] Adapted from http://en.wikipedia.org/wiki/Motor_cortex
[5] Adapted from http://www.brainm.com/software/pubs/dg/BA_10-20_ROI_Talairach/nearesteeg.htm

2.3. Evaluation

In the context of BCIs, the most common ways to evaluate and compare different designs are: user accuracy, defined as the number of "right choices" given by the classification system over the total amount of choices presented to the user; and ITR. As defined by Wolpaw et al (1998) [50], for a trial with N possible symbols in which each symbol is equally probable, the probability (P) that the symbol will be selected, within a given time (T), is the same for each symbol, and each error has the same probability. Thus, the ITR or bit rate can be calculated as follows:

$$\text{Bit Rate} = \text{Bits per symbol} \times \text{Symbols per minute} \qquad \text{Eq. 2.1}$$

Where

$$\text{Bits per symbol} = \log_2 \mathrm{N} + \mathrm{P} \cdot \log_2 \mathrm{P} + (1 - \mathrm{P}) \cdot \log_2 \left(\frac{1 - \mathrm{P}}{\mathrm{N} - 1}\right) \qquad \text{Eq. 2.2}$$

And

$$\text{Symbols per minute} = \frac{60}{\mathrm{T}} \qquad \text{Eq. 2.3}$$

A current problem with ITR is that different papers report ITR differently, or even incorrectly. Some articles calculate the ITR based on offline analysis, which is considered to have little bearing on its online performance in a field setting [51].

So, in order to apply Eq. 2.1 to evaluate a BCI system, there are some preconditions that must be met [51]:

- The BCI system must be a memory-less and stable discrete transmission channel;
- All the output commands must be equally likely to be selected;
- The classification accuracy must be the same for all target symbols;
- The classification error must be equally distributed among all the remaining symbols.

If one or more of these conditions are not met, the usage of Eq. 2.1, and thus the evaluation of the BCI system, may not be adequate and can produce an incorrect measure of ITRs. An example of BCI systems that fail to meet one or more of these conditions are the EEG-Based asynchronous BCIs. In synchronous BCIs, the system provides the user with information about when to perform an action. On the other hand, asynchronous BCIs allow users to choose when they wish to control the BCI.

This particular mechanic allows the user to remain in an idle state; any message sent during an idle period is a false positive, and should be avoided. This constraint can lead to the probability of choosing an idle state being different from the probability of selecting specific commands.

In order to standardize ITR calculations, Yuan et al. [51] suggest that:

- When reporting the ITR, N, P and T, respectively the total number of symbols, the target identification accuracy and the time needed to output a symbol, should be explicitly identified;
- To ensure an accurate estimation of classifier accuracy, enough test trials are needed. Hence, when the ITR is reported, the number of test trials should also be reported;
- Authors should include an ITR estimation that does not include error correction or other methods to increase effective throughput. If a system does employ error correction, authors should adequately describe the methods and results and, if desired, include a modified ITR as well;
- To ensure that each input symbol is equally likely to be selected, BCIs should ideally be tested with randomly generated symbols from all N symbols;
- When reporting ITRs, authors should explain all of the factors in the ITR calculation, such as whether $t1$[6] is included in the calculations;
- N should remain constant throughout the whole test;
- Results should be presented from each subject tested, including individual ITRs and statistical results. If any data were rejected from further analysis, the amount of data and the reason(s) for rejection should be described. If results are presented from subject(s) who were exceptional, this fact should be noted.

2.4. Summary

In this chapter we introduced the concept of Electroencephalography, a non-invasive method for measuring brain activity on a subject scalp, and approached a portable commercially available solution for measuring EEG activity, namely the dry electrode-based headset, Emotiv. We also described the commonly used modules in BCI systems, and briefly explained the main four paradigms associated with an EEG-based BCI: Sensory Evoked Potentials (Exogenous); Event Related Potentials (Exogenous); Event Related De/Synchronization (Endogenous); Slow Cortical Potentials (Endogenous). Due to the characteristics of our work, discussed in the following chapters, it is worth mentioning that the paradigms used and that will be further discussed in detail are those dependant on Exogenous stimuli.

Lastly, in the evaluation section we discussed the main quantitative method used for assessing the "quality" of a BCI system, namely ITR, and we also touched on the main preconditions that a given system must met in order to be considered legible for assessment.

[6] t1 is the time it takes the user to shift attention to a new target.

3. State of the art

In this chapter we focus on the relevant and recent work that has been developed on the usage of BCIs for control of interfaces using exogenous based paradigms, i.e. SSVEP and P300-based BCIs.

Firstly, we introduce the SSVEP and P300-Based BCI works, focusing on feature extraction and classification methods, common applications and stimuli systems, and lastly we discuss these works, summarizing and comparing them to each other, emphasizing their main differences.

3.1. SSVEP-Based BCIs

In this section we discuss the work that has been done on the topic of SSVEP-Based BCIs, focusing on methods used to detect SSVEP responses in the EEG, stimuli delivery systems, and commonly used interfaces for evaluation.

3.1.1. Feature Extraction and Classification

Following the presentation of visual stimuli, Steady State Visually Evoked Potentials (SSVEPs) can be recorded in the visual areas, more precisely near the striate cortex. SSVEPs can be defined as periodic oscillations, or evoked potentials, induced by rapidly repetitive visual stimulation [52].This means that an SSVEP is composed of a number of discrete frequency components, related to the frequency of the visual stimulation. Due to the particular characteristics of these signals, feature extraction and classification on SSVEP-based BCIs is mainly directed towards frequency detection methods.

The most widely used frequency detection method in SSVEP-based BCIs is the Power Spectral Density-based Analysis (PSDA). The Power Spectral Density of a signal is a representation of the power of the frequencies in a given signal. Since SSVEPs are periodic responses in the EEG signal corresponding to which visual stimulus the user is gazing at, PSDA is a straight forward method to detect the frequency, i.e. the frequency corresponding to the peak is taken as the visual stimulus frequency. This nonparametric estimation method, equivalent to the periodogram, can be calculated directly from the discrete Fourier transform (DFT) of the signal [53].

The Fourier transform is an operation that transforms data from the time domain[7] into the frequency domain. The Fourier transform of a function of time itself is a complex-valued function of frequency, whose absolute value represents the amount of that frequency present in the original function, and whose complex argument is the phase offset of the basic sinusoid in that frequency.

[7] The Fourier transform is not limited to functions of time, but in order to have a unified language, the domain of the original function is commonly referred to as the time domain.

The Fourier transform is called the frequency domain representation of the original signal. In mathematical language, the Fourier transform pair functions can be defined as [54]:

$$F(j\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \quad \text{Eq. 3.1}$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(j\omega)e^{j\omega t}d\omega \quad \text{Eq. 3.2}$$

Where $F(j\omega)$ is the Fourier transform of the function $f(t)$, making $f(t)$ the inverse Fourier transform, and $\omega = 2\pi f$.

If the ordered pairs representing the original input function are equally spaced in their input variable (for example, equal time steps), then the Fourier transform is known as a discrete Fourier transform (DFT), which can be computed either by explicit numerical integration, by explicit evaluation of the DFT definition, or by fast Fourier transform (FFT) methods. In contrast to explicit integration of input data, use of the DFT and FFT methods produces Fourier transforms described by ordered pairs of step size equal to the reciprocal of the original sampling interval. For example, if the input data is sampled for 10s, the output of DFT and FFT methods will have a 0.1Hz frequency spacing.

The spacing between the frequencies in the output is then given by the FFT resolution:

$$\Delta R_{fft} = \frac{f_s}{N_{fft}} \quad \text{Eq. 3.3}$$

Where f_s is the sampling rate of the signal in seconds, and N_{fft} the number of points in the signal with zero-padding. Although zero-padding does not provide more frequency information, it can be beneficial to separate frequencies in the output and avoid superposition.

Every signal has a frequency resolution which is the minimum spacing between two frequencies that can be resolved. The frequency resolution is determined solely by the time window or epoch, and is given by:

$$\Delta R_f = \frac{1}{T} = \frac{f_s}{N} \quad \text{Eq. 3.4}$$

Where T is the time in seconds and N is the number of points in the signal without zero-padding [54].

This means that the signal's sample rate determines the highest frequency (Nyquist frequency) that can be represented in the frequency domain and the epoch length determines the precision. For example, a 500-point epoch sampled at 1000Hz represents 0.5s and the spectral resolution is the inverse, i.e. 2Hz. If we take a 1000-point epoch sampled at 200Hz, the entire epoch is 5s long, and the spectral resolution thus becomes 0.2Hz. Since the full spectrum contains redundant information[8], typically only up to half of the spectrum is considered for both display and computational purposes. Given a real-valued time series, we can establish that the power and amplitude spectra will be even functions. Therefore, one half of the function represents all information. For example, if a time series is sampled at 1000Hz and the FFT is calculated over 512 (2^9) points of the time series, the frequency axis range covering half the spectrum is:

$$0 \rightarrow 1000 \times \left(\frac{{}^{512}/_{2} - 1}{512}\right) Hz \approx \left({}^{f_s}/_{2}\right) \qquad \text{Eq. 3.5}$$

The 1 is subtracted from ${}^{512}/_{2}$ because the frequency axis must begin at 0 (representing the DC component in the time domain). The step size on the frequency axis is $1000 \times {}^{1}/_{512}$ Hz. In a more general approach, Eq. 3.5 can be rewritten as:

$$0 \rightarrow f_s \times \left(\frac{{}^{N_{fft}}/_{2} - 1}{N_{fft}}\right) \qquad \text{Eq. 3.6}$$

The raw complex-valued output of an FFT is difficult to interpret directly. Therefore, the most common approach is to present the Power Spectrum ('S') of the given signal, where the power for each frequency is plotted along the frequency axis. This result is obtained by multiplying the FFT output ('X') with its complex conjugate ('X^*'), and normalized[9] by dividing by the number of data points (N_{fft}):

$$S = \frac{X \cdot X^*}{N_{fft}} = \frac{|X|^2}{N_{fft}} \qquad \text{Eq. 3.7}$$

The work of Lalor et al [2], in 2005, is an example of a work that uses this method. Their EEG recording system consisted of two Ag-AgCl scalp electrodes placed at sites O1 and O2. Each channel was amplified, 50Hz line filtered, and bandpass filtered over the range 0.01–100Hz by rack amplifiers, and then digitized at a sampling frequency of 256Hz.

[8] The real part of X is symmetric, corresponding to the even property of the cosine wave, whereas the imaginary part of X gains symmetry from the odd property of the sine wave.
[9] Normalization ensures that the energy (sum of squares) of the time series equals the sum of the elements in the power spectrum.

For the feature extraction system, they used $25s$ trials, divided into 50 overlapping segments. These trials were processed using two distinct methods and the segments were furthermore divided into $1s$ and $2s$ segments, for comparison on the methods. The first method used consisted of extracting each $1s$ or $2s$ segment using a Hamming window, zero-padded to 1024 samples ($4s$), and then applying and squaring the FFT. A single feature was then extracted for each segment using the following equation:

$$F_1(n) = \log\left(\frac{X_n(f_1)}{X_n(f_2)}\right) \quad \text{Eq. 3.8}$$

Where

$$X_n = mean^2\left(FFT\left(x_n(t)\right)|_{O1}, FFT\left(x_n(t)\right)|_{O2}\right) \quad \text{Eq. 3.9}$$

That is, X_n represents the square of the FFT averaged over electrode sites $O1$ and $O2$, of the nth segment $x_n(t)$, and f_1, f_2 the chosen stimuli frequencies ($17Hz$ and $20Hz$, respectively).

The second method consisted of calculating the PSD (Power Spectral Density), through FFT autocorrelation. They calculated the autocorrelation function for each segment, followed by the FFT, extracting a single feature:

$$F_2(n) = \log\left(\frac{Y_{n(f_1)}}{Y_{n(f_2)}}\right) \quad \text{Eq. 3.10}$$

Where

$$Y_n = mean(FFT(R_{xx}^n)|_{O1}, FFT(R_{xx}^n)|_{O2}) \quad \text{Eq. 3.11}$$

And the autocorrelation function R_{xx} of the nth segment $x_n(t)$ is given by:

$$R_{xx}^n(t) = E\{x_n(t_0) \cdot x_n(t_0 - t)\} \quad \text{Eq. 3.12}$$

Finally, they used a Linear Discriminant Analysis (LDA) classifier and evaluated the whole system with two subject groups, one for each method, composed of six male subjects aged between 24 and 34 with normal or corrected-to normal vision.

They measured the accuracy of the users (i.e. if the users were able to select the right decision), using the MindBalance interface (described in detail below on Section 3.1.3) and a Bit Rate analysis, where they calculated an approximate ITR of 10.3bits/min with a peak accuracy of 89.5% using a 3-second window, 15.5bits/min using a 2s window and 16.6bits/min for the 1s window. In both online and offline testing's, classification based on 2s windows exceeded that of 1s windows for all features. This was expected as a 2s window gives higher-frequency resolution and allows more accurate extraction of the SSVEP peak amplitudes.

Another work in the area of SSVEP-based BCIs using PSDA is the one develop by Vilic et al [55], in 2013. Although their system was developed for an SSVEP-based spelling interface, their feature extraction and classification method can be generally used with other applications in mind. Their recording system consisted of three gold plated electrodes placed along the test subject's scalp, a ground electrode at FPz, a reference electrode at Fz and a signal electrode at Oz. The impedances were kept around 5kΩ or lower, they used an amplifier from g.tec (g.USBamp) set to a sampling rate of 512Hz and an analog band pass filter from 5Hz to 30Hz.

Regarding their feature extraction and classification method, they applied a simple FFT (with a resolution of 0.1Hz) to two sets of data: SData, the most recent two seconds of EEG, to which is applied autocorrelation; and CData, a concatenation of up to three most recent sets of SData. Then they generate a set of classes, each representing a target frequency (i.e stimuli frequency), in which the value of each class ('C_x') is the sum of power amplitudes ('Y') from the FFT, around the relevant frequencies:

$$C_x = \sum_{H0-0,1}^{H0+0,1} |Y| + \sum_{H1-0,1}^{H1+0,1} |Y| \qquad \text{Eq. 3.13}$$

Where H0 is the fundamental frequency presented, and H1 is its first harmonic. The values in all classes are then normalized in respect to each other. The dominating class will have a value of one, but the classification only happens if at least one of three quality tests is satisfied:

- The second greatest value in SData < 0.35;
- The second greatest value in CData < 0.45;
- The same class is dominating in four consecutive iterations.

These thresholds were determined through empirical testing, in which increasing the thresholds would improve selection times for some users but at the same time reduce accuracy for others, so these values were a midterm compromise between selection times and accuracy.

Nine healthy subjects (Six males and three females, age 26.8 ± 5) participated in writing full sentences on a speller BCI interface, with 8 flickering targets and dictionary support, and obtained an ITR of 21.94 ± 15.63 bits/min and an accuracy of $90.81 \pm 4.11\%$.

Although FFT-based PSDA works well in BCI systems, it has its drawbacks. PSDA might still be sensitive to noise if the signal to be analysed is from a single channel (or a bipolar montage). Array signal processing, such as Canonical Correlation Analysis (CCA) using channel covariance information, may improve the signal-to-noise ratio (SNR) [52].

CCA can be used as an approach to analyse the frequency information in an EEG signal, and to find which frequency, among a set of frequencies, is the most prominent in the signal. CCA was first introduced for multi-channel SSVEP detection in 2007 by Lin et al [52] and their work was later complemented by Bin et al [56], in 2009.

Mathematically, CCA is a multivariate statistical method used when there are two sets of data which may have some underlying correlation [28]. The two sets of data are put in multidimensional random variables X and Y, which then suffer a few linear combinations:

$$X' = X^T W_x \, , Y' = Y^T W_y \qquad \text{Eq. 3.14}$$

Where X' and Y' are projections onto W_x and W_y, and are called canonical variants. CCA finds the weight vectors W_x and W_y which maximize the correlation between X' and Y'. According to Bin et al [56], this is done by solving the following equations:

$$\max_{W_x, W_y} \rho(X', Y') = \frac{E[X'^T Y']}{\sqrt{E[X'^T X']E[Y'^T Y']}} = \frac{E\left[W_x^T X' Y'^T W_y\right]}{\sqrt{E\left[W_x^T X' X'^T W_y\right] E\left[W_x^T Y' Y'^T W_y\right]}} \qquad \text{Eq. 3.15}$$

Solving this problem give the weight vectors W_x and W_y, but also several correlation coefficients.

When using CCA for frequency recognition in an SSVEP-based BCI, X is the multichannel EEG signal and Y is a reference signal. For each stimulus frequency, f_i, a reference signal Y_{f_i} is created. CCA will use Y_{f_i} , compare it to X and output a ρ_i which indicates the correlation between the two sets of variables (Figure 3.1). The classification can then be done simply by choosing the highest value of ρ_i, to which corresponds the target frequency f_t. Formally, this process can be written as:

$$f_t = arg \max_i \rho_i \qquad \text{Eq. 3.16}$$

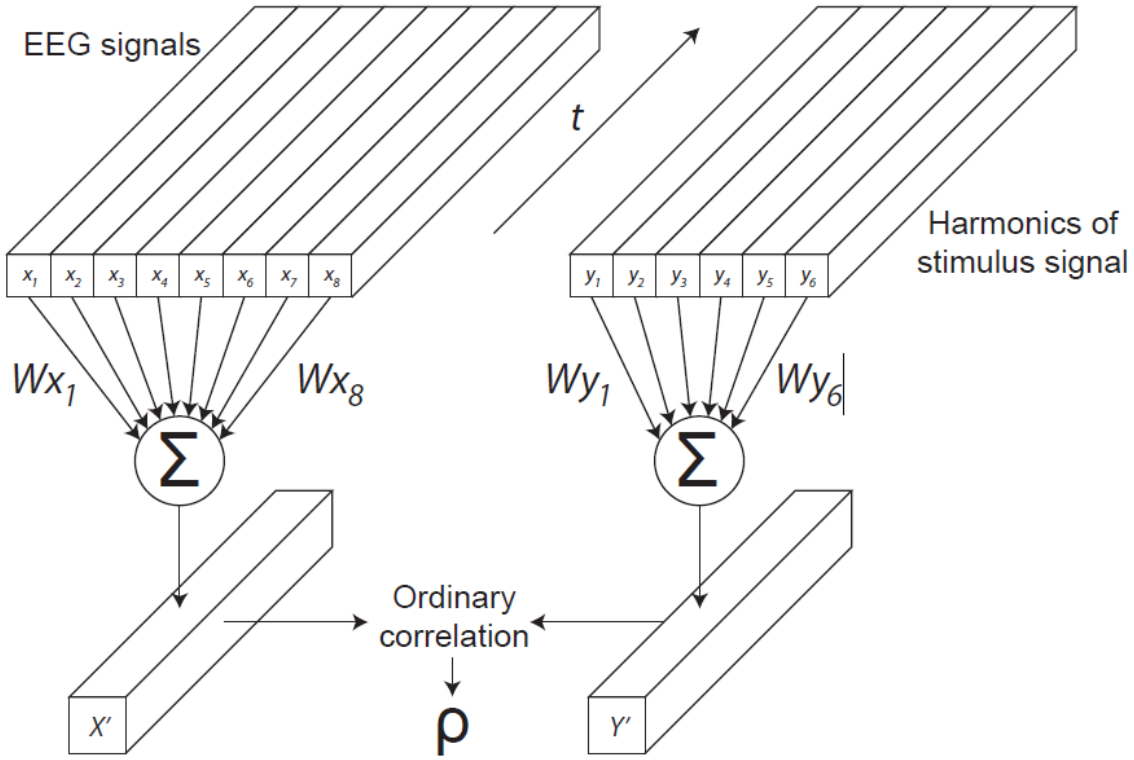


Figure 3.1 – Schematic representation of the CCA method applied to an SSVEP-based system. $x_1, x_2, \dots, x_8$ are signals from 8 distinct EEG channels, whereas $y_1, y_2, \dots, y_6$ are the sine and cosine terms created from the harmonics of the stimulus frequencies. CCA finds the weight vectors W_x and W_y which maximize the correlation between X' and Y'. Finally, ρ is the correlation between X' and Y'. Adapted from Lin et al [52].

The reference signal Y_{f_i} is a matrix created from a stimulus frequency f_i. This stimulus frequency can be decomposed into the Fourier series of its harmonics [52]. The reference signal Y_{f_i} can then be written as:

$$Y_{f_i} = \begin{pmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \sin(4\pi f_i t) \\ \cos(4\pi f_i t) \\ \vdots \\ \sin(2\pi N_h f_i t) \\ \cos(2\pi N_h f_i t) \end{pmatrix}, t = \frac{1}{S}, \frac{2}{S}, \dots, \frac{T}{S} \qquad \text{Eq. 3.17}$$

Where N_h is the number of harmonics used, T is the number of sampling points and S is the sampling rate of the EEG signal.

Cao et al [57], in 2011, presented an SSVEP-based interface speller based on the work of Lin et al. Again, as it happened before, it is worth of mention that although their system was developed for an SSVEP-based spelling interface, their feature extraction and classification method can be generally used with other applications in mind. On a 60Hz monitor, the speller system enabled the user to choose among 42 characters. The frequencies of the 16 targets started from 8Hz to 15.5Hz with an interval of 0.5Hz. The EEG signals were acquired with a g.USBamp amplifier from g.tec with a sampling rate of 256Hz. Six electrodes placed on POz, P3, P4, Oz, O1, and O2 were used as input channels. The signals were filtered by a 0.5Hz to 60Hz band-pass filter.

The system operates with a gazing interval of 2s, meaning that the user has to look at a target for 2s for the system to predict an outcome. The following second is used for the user to shift his/her gaze to the next target. The window size is set to 1s, and the correlation coefficients are calculated every 0.2s. Throughout the gazing interval, the CCA algorithm runs 5 times producing 5 correlation coefficients and commands. At least 2 out of the 5 predicted commands, C, have to be the same. If this was not the case, no command is outputted from the system, and new EEG data has to be collected. If at least 2 commands are the same, the mean is calculated of the corresponding coefficients. If the mean was larger than the empirical threshold of 0.2, C is outputted as the prediction of the target. This system achieved an average accuracy of $98.78 \pm 1.62\%$ and an average ITR of 61.64 ± 3.61 bits/min, over 4 total tested users.

3.1.2. Stimuli

Visual stimulators in SSVEP-based BCIs are typically presented to users on a computer monitor or through LEDs. On a computer monitor, commands are represented by flickering targets with its own unique stimulus sequence. These targets are commonly presented as square boxes that alternate between two colours to create the flicker. In an SSVEP BCI, each target flickers with a unique frequency and a flicker consists of an on/off cycle; a 10Hz flicker consists of 10 on/off cycles per second. When representing these targets on a computer monitor, the range of available frequencies is limited by the computer monitor refresh rate. As an example, consider a monitor with a 60Hz refresh rate, meaning it can draw a new frame 60 times per second. A 10Hz flicker cycle would consist of six frames total, alternating every three frames between on and off. This can be calculated by:

$$A = \frac{R}{f \times 2} \qquad \text{Eq. 3.18}$$

Where A is the alternating rate in frames, R is the monitor refresh rate and f is the frequency. Now consider an 11Hz flicker cycle. Representing an 11Hz flicker cycle on a 60Hz computer monitor is mathematically unfeasible since it would require the computer monitor to alternate between on and off every 2.73 frames, according to Eq. 3.18. This limits the possible frequencies on a 60Hz monitor to 1Hz, 2Hz, 3Hz, 4Hz, 5Hz, 6Hz, 6.67Hz, 7.5Hz, 8.57Hz, 10Hz, 12Hz, 15Hz, 20Hz and 30Hz.

This can, however, be avoid by controlling the stimuli using a square wave. This method was first developed and tested by Wang et al. [59], in 2010, and the idea was to approximate frequencies by using a varying number of frames in each cycle, to represent frequencies that otherwise could not be represented. A flicker at frequency f at frame i can then be calculated by:

$$flicker(f, i) = square\left(2\pi f \frac{i}{R}\right) \quad \text{Eq. 3.19}$$

Where, $square(2\pi ft)$ generates a square wave with frequency f, i is the frame index and R is the refresh rate of the monitor.

The lower frequencies ($< 5Hz$) are not practical to use in an SSVEP paradigm due to noise and high response time. Some combinations of frequencies should also be avoided in an SSVEP-based BCI since the fundamental frequency can elicit even stronger responses in the second and even third harmonic. For example, since 15Hz is the second harmonic of 7.5Hz it would be inadvisable to use both of these frequencies for flickers in an SSVEP system.

Cao et al [60], in 2012, conducted an experiment on how colours affect the performance of an SSVEP system. The theoretical background for the experiment has basis in how the human eye perceives colour and light. The human eye contains two types of photoreceptors, namely rods and cones. The rods are extremely sensitive to light, but cannot differentiate between colours. Cones are less sensitive to light, and split into three different types: those that perceive red, those that perceive green and those that perceive blue. Experiments were run for five different colours: white (255, 255, 255), grey (128, 128, 128), red (255, 0, 0), green (0, 255, 0) and blue (0, 0, 255). The numbers in the parentheses correspond to the Red Green Blue (RGB) values for that colour. CCA was used as the classification method. The article found that the correlation coefficients were higher for white and grey colours, followed by red, green and finally blue. Cao et al. conclude that white and grey outperform the other colours because they can simultaneously elicit all three cone types (red, green and blue), whereas the other colours only elicit one type of cone. This results in a more intense SSVEP response.

Finally, it is also worth mentioning that in the work of Vilic et al [55], their interface had a geometry such that, at any given time, there were either eight or six active flickering targets including the switch target. Since the size of each target was only 2.89cm^2, it would barely cover the fovea and the distance between any two targets was at least 1.7cm in any direction, so that at any point, fovea would only cover one target.

3.1.3. Applications

The first BCI-based interactive interface was created by Vidal et al, in 1977 [12]. In this interface, the user could move in four directions in a maze, by fixating on one of four fixation points displayed off-screen. A diamond-shaped checkerboard was periodically flashed between the four points, resulting in neural activity on different sites of the primary visual cortex (Figure 3.2). Using an online classification method, the VEPs were recognized, and used to move in the maze.

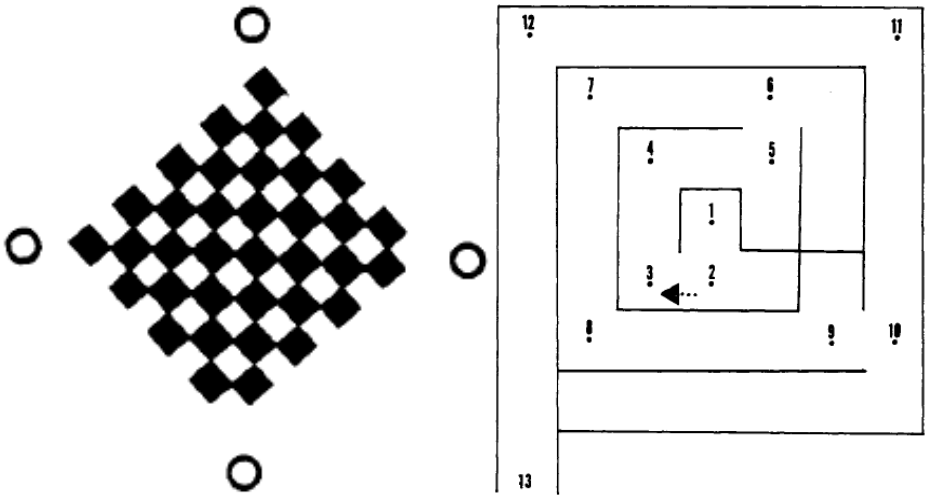


Figure 3.2 – Images that represent the interface created by Vidal et al, in 1977. The image on the left is a representation of the stimuli delivery system, were the checkerboard pattern flickers using a xenon flash, to provide visual stimulation. The four dots at the corner of the diamond represent the fixation points. For each point the stimulus target "lands" on a different retinal position with respect to the user's fovea and thus distributes neural activity on different sites of the primary visual cortex. The signal is then classified and used to move a mobile inside a maze, represented on the image on the right. Adapted from [12].

Figure 3.3 - Interface of the MindBalance game, developed by Lalor et al in 2005.

Lalor et al [2], in 2005, introduced an SSVEP-based interactive interface, named MindBalance. The SSVEPs were evoked by two different checkerboards present on each side of an avatar (Figure 3.3), phase-reversing at 17Hz and 20Hz, respectively. The goal for the user was to focus attention on one of the checkerboards to balance the avatar on a cord, inside a 3D environment. Although the environment was set in a 3D perspective, the user only had control over the 1D plain (controlling the left/right balance of the avatar). In the course of the game, the avatar walks toward the user, and stumbles every $1.5 - 5.0$s to one side chosen randomly. The user must then direct his gaze and focus on the checkerboard on the opposite side of the screen, to restore the balance of the avatar. The avatar remains in an "unbalanced" position for 3 seconds, to allow for the user to react and make a decision.

They used a setup with two machines, one for the rendering of the 3D interface and another for the real-time acquisition and processing of the EEG data, using a combined graphics, signal processing, and network communications engine. Whenever a decision on the fate of the avatar was required (after the 3s interval), a request was made to the signal processing machine which sent back a decision based on the most recent feature extracted from the EEG.

Chumerin et al, in 2013 [58], developed an SSVEP-based interface with a similar concept to the one developed by Vidal et al. The user controls an avatar (depicted as Homer Simpson's head) in a maze-like environment (Figure 3.4). The task was to navigate the avatar to the target (i.e. a donut) through the maze. The maze consisted of square cells surrounded by walls through which the avatar could not pass. The default mode of the interface was the free-play mode (Figure 3.4**a**), where each maze was generated randomly and the avatar's movement is restricted only by the maze walls. Advancing from level to level, the size of a maze cell would decrease, allowing more complicated mazes to fit into the same (fixed-size) screen region designated for the maze. For testing purposes, they also used a special experimental mode with a predefined maze, where the avatar was allowed to move in predefined directions only (Figure 3.4**b**).

The relative speed of the avatar, defined as the number of cells the avatar can visit per second, was predefined and did not depend on the level of the interface. The user was able to control the avatar by gazing at the flickering arrows (showing the direction of the avatar's next move) placed in the periphery of the maze. Each arrow would be flickering with its own unique frequency, which was predefined (default settings) or set accordingly to the player's preference. The interface was implemented in MATLAB as a client–server application, using the Psychophysics Toolbox extensions.

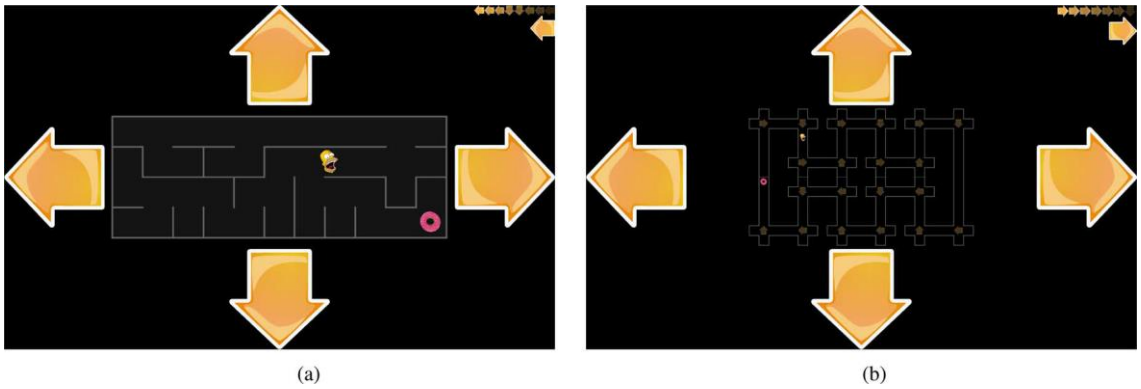


Figure 3.4 – Screenshots of "The maze" interface created by Chumerin et al, in the (a) free-play and (b) experimental modes. The decision queue can be seen in the upper-right corner of the screen, and the "final" command is shown just below the queue. The stimuli used to control the avatar is represented as four big arrows on the periphery of the maze. Adapted from [58].

3.2. P300-Based BCIs

In this section we discuss the work that has been done on the topic of P300-Based BCIs, focusing on methods used to detect P300 responses in the EEG, stimuli delivery systems, and commonly used interfaces for evaluation.

3.2.1. Feature Extraction and Classification

The signal processing of evoked potentials remains a difficult and somewhat unsolved problem. This may be due to a lack of "zeal" on the individual subject's part, normal subject signal variations, outside contamination of artifacts, or even just because of the low signal-to-noise ratio available from an EEG signal.

Not all P300 recognition techniques are suitable for use in an online BCI. Averaging over trials has been used to improve signal detection, but this technique requires a costly trade-off between recognition accuracy and the time taken to recognize a particular signal. Traditionally, P300s are obtained by averaging EEG signals from specific electrode sites over many trials. Averaging to obtain an evoked potential contains the following benefits:

- It reduces the contribution from unrelated spontaneous EEG signals;
- It reduces spurious noise;
- It allows the observation of a response that would otherwise be unobservable.

An average may tell the clinician of an abnormality that a single trial could not, because of the natural variance in both the latency and morphology of an P300. Results suggest that P300 amplitude does not stabilize until approximately 20 trials have been averaged, although peak latency does not change significantly during this time frame [61]. Even though the P300 signals may vary, the online nature of a BCI requires the P300 signals to be recognized in a timely fashion if they are to be useful. There is a distinct trade-off between the time taken to recognize a P300 signal and the percent of P300s that are recognized correctly.

The alternative to trial averaging is called Single Trial Recognition (STR). In this method a single ERP trial is taken, analysed and classified. The main downsides of this method are: it needs a faster/simpler method for feature extraction and classification; and it is more prone to errors. On the other hand, it allows for faster response times, faster computational times and it can even be used with a special technic for oddball stimuli presentation that shortens the response time even more [62].

Either we consider averaging or single trial recognition, due to the inherent characteristics of a P300 response the methods used to extract and classify features are mostly centred on time domain processing. Given that the time domain of an EEG signal is more prone to artifacts and lower SNR usually most researchers tend to combine multiple methods of feature extraction and classification, for better results. Methods for feature extraction and classification range from the heavier processing algorithms like stepwise linear discriminant analysis (SWLDA), support vector machines (SVM) and Wavelet analysis [63] to more lighter algorithms, commonly used in Single Trial Recognition systems, like Area analysis, Peak picking and Correlation analysis [62].

The Area analysis method is simply a calculation of the 'area under the curve' of the P300 component of the EEG signal between $250ms$ and $600ms$ (see Figure 3.5) [62]. If we consider this method alone, this means that the extracted feature is a value of the area under the curve of an EEG signal between 250ms and 600ms after the onset of the stimuli, and the corresponding classification would be done using the values from multiple classes, each corresponding to the different stimuli, in which the highest value would be chosen as the command class.

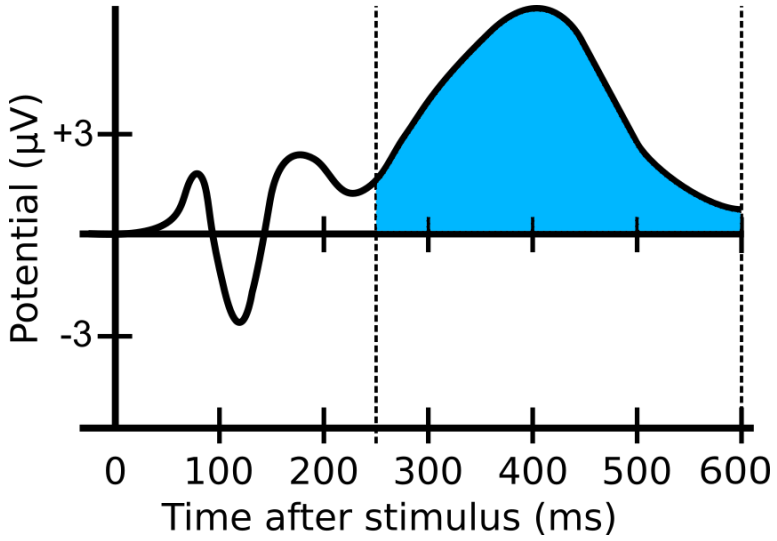


Figure 3.5 - Graphical representation of the temporal evolution of a P300 EEG signal, highlighting in light blue the zone defined for the Area analysis. As mentioned in the text, the Area analysis method uses a defined temporal interval, between 250 and 600ms, returning the total area under the curve for that interval.

But this method alone is very prone to false positives, artifacts and noise, so it is usually used in conjugation with the Peak picking method, which is on its own a better method.

Peak picking is a simple algorithm to classify a P300 component using the difference between the minimum and maximum amplitude in an epoch. Usually it works by calculating the difference between the maximum amplitude of an EEG signal between 250ms and 600ms after the onset of the stimuli and the minimum amplitude of an EEG signal between 0 and 250ms after the onset of the stimuli (see Figure 3.6).

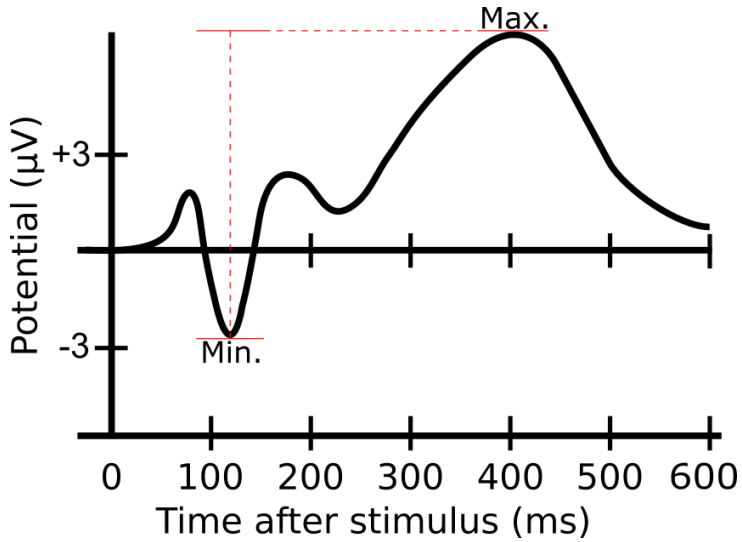


Figure 3.6 - Graphical representation of the temporal evolution of a P300 EEG signal, highlighting in red the peaks used in the Peak Picking method. As mentioned in the text, the Peak Picking method searches the curve for its lowest value in the interval between 0 and 250ms, and for its highest value between 250 and 600ms, returning the difference in amplitude between these two peaks.

As it happens with the Area method, the corresponding classification would be done using the values from multiple classes, each corresponding to the different stimuli, in which the highest value would be chosen as the command class.

A slightly more complex, but still easily calculated recognition algorithm is correlation. Correlation may be looked at as template matching when the correlation is performed between single trials and a template of what each kind of trial should look like. The biggest downside of this method is that it requires an epoch prior to the online run, which means that systems based on this method require some sort of "training" session, in which the user EEG signal is being recorded and averaged to create a template ERP signal. In the ERP context, the correlation algorithm is given by [64]:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \qquad \text{Eq. 3.20}$$

Where x represents the data for a single trial epoch, y represents the template signal, in this case a P300 signal average or a Non-P300[10] signal average, $cov(x,y)$ represents the covariance of x and y, and σ is the standard deviation of the appropriate signal, x or y. Note that the covariance is a scalar, because the x and y vectors both represent multiple instances of a single variable.

One of the first works that introduced the concept of single trial recognition, was the work develop by Farwell et al [62], in 1988. Four healthy subjects, 3 females and 1 male, 20-36 years old, participated in the study, which consisted in communicating a 5- letter word to a computer, through an ERP-based BCI interface (i.e. "BCI-speller"). The primary purpose was to determine the number of trials and the rate of event presentation that were required to achieve a specified level of accuracy in communication. The EEG was recorded from Ag-AgC1 Beckman Biopotential electrodes placed at the Pz site. The authors refer that they chose this site because in previous works it was mentioned as the site where the largest amplitude P300 was recorded in young adults. They also recorded Electrooculogram (EOG) from sub- and supraorbital electrodes (above and below the right eye). The subjects were grounded at the forehead; electrode impedance did not exceed $5\text{k}\Omega$. The EEG signal was amplified by Grass Model 12 amplifiers with low- and high-pass filters set at half-amplitude frequencies of 35Hz and 0.02Hz, respectively. These signals were digitized at a rate of 50Hz.

[10] In this context, a Non-P300 signal can be perceived an ERP signal in which the P300 peak is absent, i.e. the user was gazing at another target, in the case of a multi-target experience, or the target that the user was gazing at was not the goal target.

Subjects were presented with a 6-by-6 matrix whose cells contained the letters of the alphabet as well as several 1-word commands for controlling the system. The matrix was displayed on a computer-controlled Cathode Ray Tube (CRT). Subjects were instructed to attend to a given letter and to keep a running mental count of the number of times it flashed. In the first session the subjects completed 6 blocks of 120 trials each. The first session was concluded with a block of trials in which was used a real-time signal-detection algorithm to allow subjects to communicate the word 'BRAIN' to the computer. Subjects selected each of the letters in the word 'BRAIN' in turn, and silently counted the flashes of the row or column containing the letter until the system displayed the letter it had selected in a specified position on the screen. After the letters spelling the word 'BRAIN' had been displayed, the subject selected the 'TALK' command, and the word was sounded by means of a Votrax speech synthesizer. In a few cases the computer displayed a letter other than the one on which a subject was focusing. The subject then focused on the BKSP ('backspace') command to correct the error. Each trial contained 6 distinct events, namely, the flashes of each of the rows or columns, only one of which was task-relevant. For analytical purposes, each trial was divided into 6 data windows or subtrials, each consisting of the data for $600ms$ after onset of the flash of a row or column. For each subtrial they computed a score that measured the magnitude of the P300 in the epoch following the presentation of the row or the column.

Four different algorithms were used to compute the magnitude of the P300 in each epoch, namely: **(a)** SWLDA; **(b)** Covariance; **(c)** Peak Picking; and **(d)** Area. In short, **(a)** SWLDA yields a score that measures the 'distance' between each epoch and the mean of a group of trials known to include a P300, by applying a discriminant function to the data from each epoch. Therefore, as it happens with Covariance there is the need for a 'training' session prior to the online session, to create a P300 template. This template was developed on the basis of a 'training set' of trials recorded while the subject was focusing on the first 2 letters ('B' and 'R'). The remaining ERPs provided the 'analysis set.' They used the training set data to compute the discriminant function that distinguished between the 'attended' subtrials ($600ms$ following the flash of a row or column containing the attended cell) and the 'unattended' subtrials ($600ms$ following the flash of a row or column not containing the attended cell). The **(b)** Covariance was computed using all of the points in the $600msec$ epoch, and the P300 scores in the analysis set were derived by computing the covariance of each subtrial with the template. Regarding **(c)** Peak picking, the amplitude of the P300 was computed as the difference between the lowest negative point prior to the P300 window (defined as the time range within which the average attended wave form in the training set for each subject was positive) and the highest positive point in the P300 window. The window for the P300 ranged typically between 220 and $500ms$. Finally, the **(d)** Area of the P300 was calculated as the sum of data points in the P300 window, defined in **(c)**.

The values attained from the above analyses were then used to determine the letter upon which the subject was focusing attention.

Row and column scores given by the respective algorithms were summed to compute a unique score for each cell in each pair of trials (1 trial in which rows were flashed and 1 trial in which columns were flashed). For example, the score for 'B,' which was located in the first column and the second row, was the sum of the score for the first column and the score for the second row. The scores computed for each letter were summed across trials to determine which cell was identified as the cell selected by the subject. Each test could yield 1 correct response or 1 of 35 possible errors. The test was considered a 'hit' if the algorithm yielded the largest total score, summed across trials, for the letter on which the subject was focusing. For example, if the subject was attending to the letter 'B' and 6 trials were being considered in the analysis, a correct response would be achieved if the total of the 6 'B' scores - the scores for the rows and columns containing 'B' - was greater than the total of the 6 scores for any other cell in the matrix. SWLDA and peak picking proved to be the most efficient algorithms, which in the context and purpose of the experience means the fastest algorithms to reach both 80% and 95% accuracy in 3 out of 4 cases.

Later in 2004, Bayliss et al [64], developed an ERP-based BCI set in a Virtual Reality (VR) environment, using Peak picking and Correlation, as chosen algorithms for feature extraction and classification.

Seven electrode sites were arranged on the heads of nine subjects with a linked mastoid reference. Sites Fz, Cz, Pz, $P3$, $P4$, as well as an upper and lower vertical EOG channels were used. For online recognition and analysis, EOG artifacts were regressed out of the signals of interest using an algorithm by Semlitsch [65]. The EEG signals were amplified using Grass amplifiers with an analog bandwidth of $0.1-100\text{Hz}$. Electrode impedances were $2-10\text{k}\Omega$ for all subjects. An epoch size from -100ms (prior to stimulus onset) to 1500ms was specified for a total epoch size of 1600ms. The data was recorded continuously and saved to a file.

The experiment consisted of four tasks: Calibration; VR condition; Monitor condition; and Fixed Display condition. The calibration task was used to ‘train’ a signal processing algorithm (i.e. Correlation) on a particular subject’s P300 signal response. A total of 300 stimulus presentations were presented to each subject. In this task, subjects were told to count only the lamp sphere flashes; thus, in this task only the lamp sphere flashes were task-relevant, and these flashes should have caused a P300 response. Since the spheres flashed randomly over the five controllable items, approximately 60 lamp flashes occurred over the course of 5min. The remaining tasks were accomplished in a randomized block order and lasted for approximately 5min each (250 stimulus presentations with the sphere flashed randomly on items). These tasks involved online single trial classification of the P300 in order to control the different items in the apartment. The time taken for these trials depended on how many items were controlled, as subjects received feedback for each item for which the signal classification algorithm classified the trial as a P300 component trial. During each task, the goal was chosen randomly, and the subject tried to achieve the goal for up to 50 presentations of the goal stimulus.

When the goal was achieved, an action involving visual feedback occurred in the virtual apartment (e.g., the room was lightened when the light was turned on). During the waiting period for this visual feedback (1.5 sec), no new stimuli were presented. Then, the next goal was randomly chosen.

While the experiment involved online classification and feedback, an offline analysis was done to compare the obtained P300 components between different conditions. Only epochs with a maximum vertical EOG signal of less than $50\mu V$ were used. This reduced the possibility of EOG contamination of the averages. For analysis, the epochs corresponding to the EEG signals from the VR experiment were divided into goal and non-goal responses. For example, when lamp lighting was the goal, it was a goal response, while turning on the TV when the lamp was a goal was a non-goal response. These were further divided into responses with P300s present and P300s absent, yielding four categories of responses: goal P300 present; goal P300 absent; non-goal P300 present; and non-goal P300 absent. The Peak picking algorithm used was the following:

$$\max(x) - \min(x) \geq \tau \begin{cases} 1 \; if \; P300 \; Present \\ 0 \; if \; P300 \; Absent \end{cases} \quad \text{Eq. 3.21}$$

Where x is a vector that represents the data for a single EEG response and τ represents a threshold required to accept a P300. For recognition, the time window with the best result was $300-600\text{ms}$, and the voltage difference threshold was varied in experiments to yield the best result for each user.

Regarding Covariance analysis, the algorithm used is the one present in Eq. 3.20, and to determine if an EEG response contains a P300, the correlation between the EEG response epoch and the P300 average was compared with the EEG response epoch and non-P300 average according to the following algorithm:

$$\begin{aligned} &if \; (\rho_{P300} > \tau \; and \; \rho_{P300} \geq \rho_{nonP300}) \\ &then \; P300 \; is \; Present, else \; P300 \; is \; absent \end{aligned} \quad \text{Eq. 3.22}$$

Where ρ_{P300} is the correlation between the EEG response and P300 average, $\rho_{nonP300}$ is the correlation between the EEG response and non-P300 average, and τ is the threshold set to determine the extent of the desired correlation. This threshold was also varied in experiments to yield the best results. If the EEG response did not correlate with a P300 or non-P300 average the algorithm is indeterminate. During experimentation, all indeterminate responses were treated as not having P300s.

Peak picking and correlation achieved the same overall VR accuracy, 81%; however, correlation more significantly impacted the subject's accuracies ($p < 0.008$), than peak picking ($p < 0.04$, p values were derived using the Wilcoxon signed rank test).

3.2.2. Stimuli

As mentioned before, the P300 often is elicited with a two stimulus discrimination task. This procedure has been dubbed the 'oddball paradigm', since two stimuli are presented in a random series such that one of them occurs relatively infrequently, i.e. the oddball. The presentation of a stimulus in an oddball paradigm can produce a positive peak in the EEG, $300ms$ after onset of the stimulus, which can be either visual, auditory or somatosensory [43]. This type of stimuli was introduced in 1975 by Squires et al [71], and is represented in Figure 3.7. Although the oddball task has been used most often in basic ERP studies, the P300 potentials can also be obtained from a 'single-stimulus' paradigm, by replacing the standard stimuli (non-oddball) with 'silence' [44].

As mentioned before, the online nature of a BCI requires P300 signals to be recognized in a timely fashion if they are to be useful. There is a distinct trade-off between the time taken to recognize a P300 signal and the percent of P300s that are recognized correctly. In order to improve upon systems that use P300-based Single Trial Recognition methods, Farwell et al [62] introduced a stimuli system in which the Inter Stimulus Interval (ISI) was varied in such a way that the epoch from a previous stimuli would overlap with the epoch from the next stimuli (Figure 3.8).

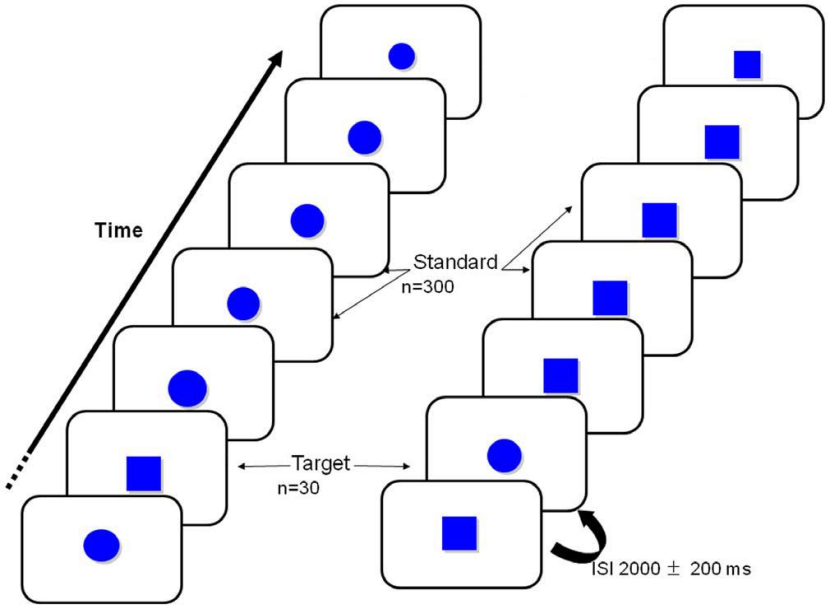


Figure 3.7 – Schematic representation of two visual oddball paradigm experiments. The oddball ($n = 30 \cong 10\%$) stimuli is a square, on the left experiment, and a circle, on the right experiment. The Inter Stimulus Interval (ISI) for both experiments was $2000 \pm 200ms$. Adapted and modified from the work of Bocquillon et al [72].

More specifically, in every trial each of the 6 rows of the matrix from the interface, or each of the 6 columns, was intensified for a period of 100ms. In the first session the ISI was 500ms. In the second session data was acquired with both a 500ms and a 125ms ISI. Half of the blocks were run with a 125ms ISI between the onset of the intensification of a given row or column and the onset of the intensification of the next row or column to be flashed and half with a 500ms ISI. The ISI was measured from the beginning of the intensification of each row/column and of the subsequent row/column to be intensified. The rows were selected for intensification in a random order, and then the columns were intensified in a similar manner.

The relevant data consisted of the EEG digitized for 600ms after the onset of each flash. The EEG was digitized continuously from 20ms prior to the first flash in each trial to 600ms after the sixth flash. The subsequent trial began approximately 620ms after the sixth flash. The inter-trial intervals (ITIs) measured from the beginning of one trial to the beginning of the next trial, then, were 1245 and 3120ms for the 125 and 500ms ISIs respectively.

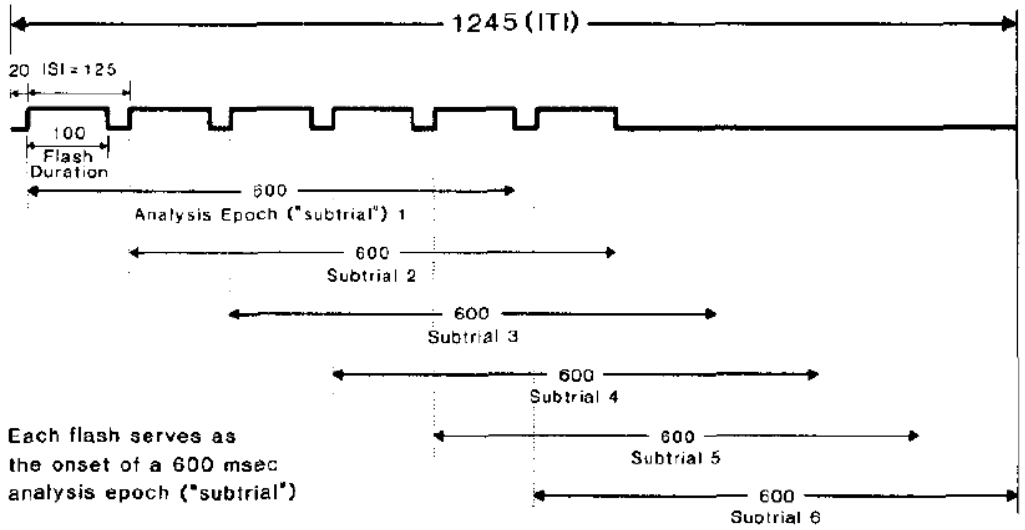


Figure 3.8 – Schematic representation of the stimuli system developed by Farwell et al [62], for a 125ms ISI. The EEG recording began 20ms prior to the onset of the first stimuli. Each stimulus lasted for 100ms, and the resting time between stimulus was 25ms. Here we can also see that, due to the small ISI, the various epochs overlap, creating an ITI shorter than what would be expected if the stimulus was only delivered after the 600ms epoch.

This technique allows for shorter ITIs and for faster response time, because it allows for one P300 to be elicited during the recording of the last P300. This causes a 'benign' superposition, due to the P300 wave characteristics, i.e. the P300 appears 300ms after the onset of the stimuli, and with this technique the various P300s will not superpose.

Although the Emotiv headset has been integrated in a vast number of BCI applications, its use to obtain P300 signals generates some concerns among researchers. The technical report presented by H. Ekanayake [73] in 2010, seeks to respond to these concerns by performing an empirical evaluation of the Emotiv headset, using the popular P300-speller interface created by Farwell et al [62]. Since the main concerns revolve around the fact that the Emotiv headset lacks the commonly used electrodes for a P300 signal recording, e.g. $Fz, FCz, Cz, Cpz, Pz, POz,$ etc…, the researcher decided to record using every channel in the headset. He recorded the signals throughout multiple sessions, averaging the epochs to obtain a clear P300 signal. According to the author results, the P300 response can be seen, with varying levels of clarity, at locations $P7, P8, O1, O2, FC5, FC6, F3$ and $F4$, with no ERP like waveform visible at other locations.

3.2.3. Applications

Curtin et al [66], in 2012 developed a P300 oddball paradigm based BCI for spatial navigation control in a virtual environment. To achieve this, they used a set of distinct tools, namely: MazeSuite® [67], [68], a free collection of software that allows to easily create and visualize 3D environments, as a virtual environment system; and BCI2000 [69], a general-purpose system commonly used in BCI research, as a stimuli delivery system. The whole system, presented below in Figure 3.9, consisted of two independent PCs, one used for signal acquisition, processing and stimuli presentation (Acquisition PC), and the other solely for environment presentation and feedback (Visualization PC). The communication between PCs was done using a TCP/IP component of MazeSuite®.

EEG measurements were taken from 9 electrode sites (FCz, Cz, $CP3$, CPz, $CP4$, $P3$, Pz, $P4$, Oz) with $A1$ and $A2$ serving as reference and ground locations. Recordings were taken using a sampling frequency of 1kHz with a software band pass filter between 0.5Hz and 60Hz and an additional notch filter at 60Hz to remove electrical line noise. Measurements were recorded by Acquire and transferred to BCI2000 using the NeuroscanAccess module, and finally to a custom analysis module in Matlab for online processing.

Subjects were asked to gaze at the target icon and count the number of times the target icon flashed. After each set of flashes, the recording was processed by Matlab using SWLDA, and the resulting classification was forwarded to MazeSuite for execution of the virtual movement command, i.e. movement of the avatar. The subject attention was modified using a series of dialogues that would direct the user to attend either the stimuli matrix (before the P300 stimulus) or the virtual environment (shortly after the end of the P300 stimulus).

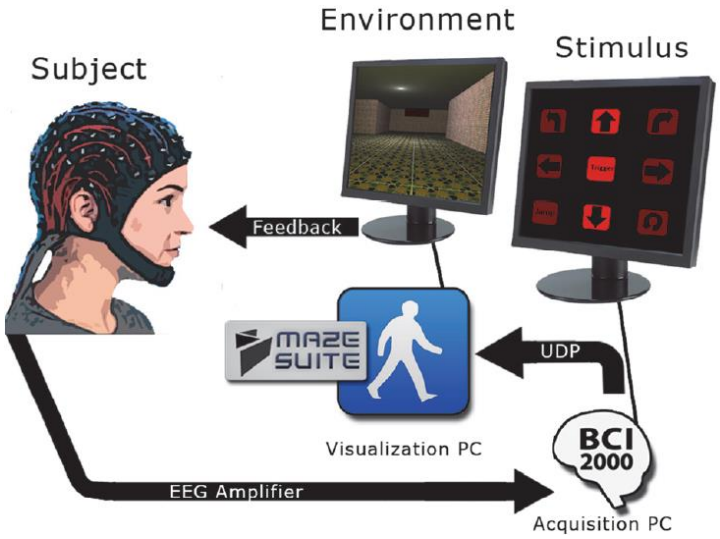


Figure 3.9 – Schematic diagram of the experimental setup from Curtin et al 2012 study [70]. The Acquisition PC was responsible to acquire and analyse the EEG data and handle the stimulus timing and presentation. Classifications based on extracted P300 components were then forwarded as command signals to the Visualization PC, and used to update the virtual environment.

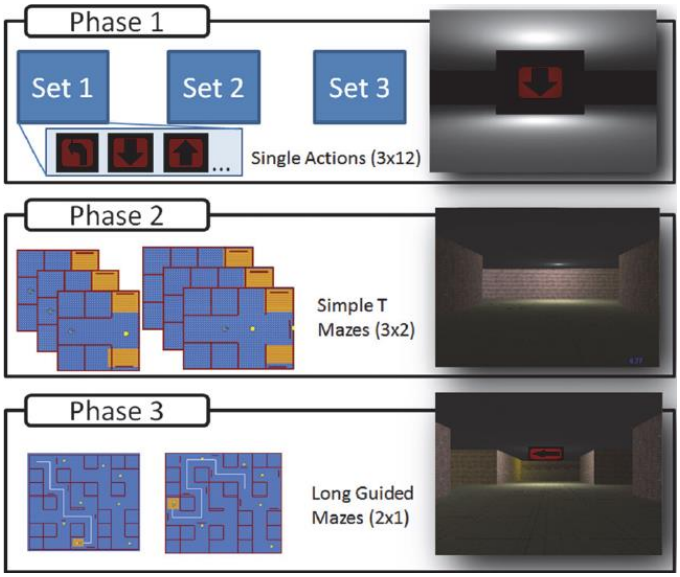


Figure 3.10 – Schematic representation of the navigation task protocol used in the study conducted by Curtin et al in 2012 [70].

In order to test the system, they used three distinct phases, which are presented in Figure 3.10. The first phase was a training segment, were the user was presented with a "single-action" environment in which the required action was clearly indicated to the subject via presentation of the target icon. The second phase presented the user with several short "T-shaped" maze environments in order to test system accuracy and introduce the user to sequential command execution for navigation. Finally, the third phase consisted of two long guided mazes with the intention of measuring the user's performance in a complete navigational task. The reported overall mean bitrate per trial was 2.06.

3.3. Discussion

There are numerous works developed on the field of BCIs. The ones presented here are merely meant to be representative of their particular topics, and were selected for being both inspirational and necessary throughout the development of our proposed solution.

Table 3.1 is a summary of these works, specifically in the topic of SSVEP-based BCIs. We chose a few characteristics that we believe are the core of these works, and differentiate them from each other. Starting from the left, the characteristics are:

- "Electrode System" - whether it is a system based on wet or dry electrodes;
- "Used Electrodes" – which electrodes, from the $10-20$ system, were used in the study;
- "Portability (device)" - if the recording system has the ability to be portable, and which recording system type was used;
- "Feature Extraction" - what is the feature extraction method used;
- "Classifier" - what type of classifier is used;
- "Harmonics" - if the harmonics of the fundamental frequencies are considered in the feature extraction system;
- "System" - if the system that is used to put everything working together is single or multiple;
- "Interface" - what type of interface is used to test the system;
- "Accuracy" – the average accuracy;
- "ITR" – the average ITR of the system;
- "#Users" – number of test subjects used to evaluate the BCI system.

Table 3.1 - Summary of the main works developed on the topic of SSVEP-based BCIs. PSDA – Power Spectrum Density Analysis; CCA – Canonical Correlation Analysis; MEC – Minimum Energy Combination; LDA – Linear Discriminant Analysis; N/A – Not Available; ITR – Information Transfer Rate.

	Lalor et al [2]	**Vilic et al [55]**	**Cao et al [57]**	**Chumerin et al [58]**
Electrode System	Wet	Wet	Wet	Dry
Used Electrodes	O1; O2	Oz	POz; P3; P4; Oz; O1; O2	N/A
Portability (device)	No (g.tec)	No (EEG system)	No (EEG system)	Yes (Emotiv)
Feature Extraction	PSDA	PSDA	CCA	MEC
Classifier	LDA	Classes	Direct	Direct
Harmonics	No	Yes	N/A	Yes
System	Multiple	Multiple	Multiple	Single
Interface	1D "Game"	Speller	Speller	Maze
Accuracy (%)	89.5	90.81	98.78	80.37
ITR (bits/min)	15.5	21.94	61.64	N/A
#Users	6	9	4	20

Starting with the first two criteria, electrode system and portability, the works of Lalor et al, Vilic et al and Cao et al use a wet electrode system, with very little portability, adequate to a research/clinical setting and not designed with a more daily use approach. Chumerin et al introduce the usage of a portable dry electrode system, using the Emotiv headset. Regarding the used electrodes, all the works used the electrodes from the Occipital area in the $10-20$ system. Since Chumerin et al reportedly rotated the Emotiv headset by $180°$ so that there were more electrodes in the Occipital area, there is no guarantee as to if the new electrode positions can be correlated to the $10-20$ system.

Moving on to the signal processing section of the works, Lalor et al and Vilic et al both use a straightforward PSDA approach with the main difference that Vilic et al use a simple classifier system based on classes which served as an example for the construction of our classifier system, and Lalor et al used a heavy LDA algorithm, and also Vilic et al used the information from more frequencies other than the fundamental harmonics and Lalor et al did not considered this information.

Regarding the feature extraction system, Cao et al used the semi-light CCA algorithm and Chumerin et al used a heavy MEC (Minimum Energy Combination) algorithm, both using a classification system based on the direct analysis of the results from their feature extractions system.

Moving on the topic of the system and interface, Lalor et al, Vilic et al and Cao et al made use of multiple and distinct systems for signal processing and stimuli delivery, which, in our opinion, hinders the operability of the interface, making it harder for the user to interact without the need for technical assistance.

Finally, we considered a few important criteria from the experimental scenario, namely ITR, Accuracy and number of users. It is worth mentioning that the values of ITR and Accuracy depicted on the table were the mean values of the experiment and should not be considered as a definitive measure of quality of the overall BCI system. With this regard, we also need to mention that the number of users that were considered in the experiment evaluation is the main factor that hinders the ability to consider ITR as a measure of quality. The only work that approaches a significant number of users is the work of Chumerin et al, with 20 users and an average accuracy of 80.37%. Unfortunately, the authors decided not to perform an ITR measure, which in our opinion is related to the fact that some of the features presented on the interface makes it impossible to be measured (See Section 2.3).

Table 3.2 – Summary of the main works developed on the topic of P300-based BCIs. SWLDA – Step-Wise Linear Discriminant Algorithm; N/A – Not Available; STR. – Single Trial Recognition; VR – Virtual Reality; ITR – Information Transfer Rate; † - In the context of the work of Farwell et al, the ITR is only measured when the user achieves an intended accuracy of 95%; ǂ - Curtin et al reported a bitrate per trial measure, which does not fit into an ITR measure.

		Farwell et al [62]	**Bayliss et al [64]**	**Curtin et al [66]**
Electrode System		Wet	Wet	Wet
Used Electrodes		Pz	Fz; Cz; Pz; P3; P4	FCz; Cz; CP3; CPz; CP4; P3; Pz; P4; Oz
Portability (device)		No (EEG System)	No (EEG System)	No (EEG System)
Feature Extraction	**Peak Picking**	Yes	Yes	No
	Area	Yes	No	No
	Correlation	Yes	Yes	No
	Other	SWLDA	N/A	SWLDA
	Combination	No	No	No
Calibration		Yes	Yes	Yes
Classifier		Direct	Direct/ Threshold	Direct
Trial		STR	STR	Average
System		Multiple	Multiple	Multiple
Interface		Speller	VR	Maze
Accuracy (%)		95†	81	88.9
ITR (bits/min)		12.0†	85	N/Aǂ
#Users		4	9	9

Table 3.2 is a summary of the works developed on the topic of P300-based BCIs. Some of the characteristics presented on this table are the same as the ones presented on the previous table, therefore we will only describe the ones that are distinct. Starting from the left, the distinct characteristics are:

- "Calibration" – if the system is based on a method that requires calibration;
- "Feature Extraction" – what is the feature extraction method used (Peak Picking, Area, Covariance, or other) and if it is a combination of multiple methods;
- "Trial" – if the system is based on trial average or Single Trial Recognition (STR);

Starting with the electrode system and portability, in a similar way to the works of SSVEP-based BCIs, the works of P300-based BCIs are predominantly based on wet electrodes, with very little portability, adequate to a research/clinical setting and not designed with a more daily use approach. As for the used electrodes, generally all works used the electrodes from the Frontal, Parietal and Central areas in the $10-20$ system, with a few non-significant differences.

Regarding the signal processing criteria, we chose feature extraction methods, classification methods and trial handling. Farwel's et al and Bayliss's et al works focused on the independent use of basic, light and straightforward algorithms (with the exception of SWLDA of Farwell's et al work), namely Peak Picking and Covariance on both, with the addition of Area and SWLDA on Farwell's et al work. Curtin's et al work focused solely on SWLDA. All of these works used a direct classifier, with some of the works using decision thresholds, based on the output from the various feature extraction methods. Finally, considering the calibration setup, all works make use of calibration to gather signal for the feature extraction methods that require a classifier (e.g. Correlation and SWLDA) and in the work of Bayliss et al in particular, calibration was also used to construct a threshold per subject for the Peak Picking method.

Moving on the topic of the system and interface, all of the mentioned works made use of multiple and distinct systems for signal processing and stimuli delivery, which as it happened with the case of the SEP works, in our opinion, hinders the operability of the interface, making it harder for the user to interact without the need for technical assistance.

Finally, and similarly to the SEP works, we considered a few important criteria from the experimental scenario, namely ITR, Accuracy and number of users. Again, it is worth mentioning that the values of ITR and Accuracy depicted on the table were the mean values of the experiment and should not be considered as a definitive measure of quality of the overall BCI system. With this regard, we also need to mention that the number of users that were considered in the experiment evaluation is the main factor that hinders the ability to consider ITR as a measure of quality. Furthermore, Curtin et al did not consider an ITR measure, but instead chose to measure bitrate per trial, which we think is related to interface design choices that limited the measurement, as it happened with the case of the Chumerin et al work on the topic of SEP-based BCIs,

3.4. Summary

In this chapter we discussed the most relevant works that have been developed over the years on the field of BCIs, namely on SEP and ERP-based BCIs, with a special focus on feature extraction methods, common interfaces and stimuli systems. We also presented an analysis and discussion on what the various works mean to our proposed solution, highlighting their main characteristics and differences.

In conclusion, and considering the context of our work, where we want to develop a "plug-and-play" BCI system, we consider that it should address the three main problems that affect BCI systems:

- **Setup:** The setup must be as fast as possible, with minimal effort and requiring no trained personnel. Therefore, we considered the usage of dry electrodes and the portability of the system, to minimize and facilitate the setup of the electrodes;
- **Calibrate:** Since calibration is counterintuitive in the context of "plug-and-play" we decided to not use it, in order to facilitate the user interaction with the interface and to evaluate the performance of the algorithms on these conditions;
- **Operate:** Although our purpose is to develop the interface to test and evaluate the designed BCI system, we should also consider that the target user for this type of interface is the common user, with no advanced knowledge of signal processing. Therefore, in this context an interface should be designed to use a single platform, in order to minimize user inputs and software setup time.

4. Developed BCI Systems

In this chapter we present and discuss the main components of the architecture of the systems that we developed in order to achieve our goal of creating a "plug and play" BCI system. Our solution offers an easy and quick setup, without training time ("calibrate") and no user intervention on configurations ("operate").

To achieve this, we developed two distinct systems based upon the two exogenous paradigms – i.e. SSVEP and P300. The two solutions described here were then used to identify the best exogenous paradigm for a "plug and play" BCI system. To compare these two systems, and test/evaluate the BCI system, we used them to control a BCI-based interactive interface using only inputs from EEG signals recorded with the Emotiv EPOC Headset.

The interactive interface consisted of a $2D$ top view maze, in which the BCI system controls the movement in four distinct directions of an avatar inside the maze. These systems were developed and optimized to identify four commands, corresponding to each of the four possible movement directions of the interface. To be able to recognize more commands, the system configurations would have to be reviewed.

In the following subsections, we present and discuss the architecture of the developed solutions, focusing on the main modules of each system.

4.1. BCI System General Architecture

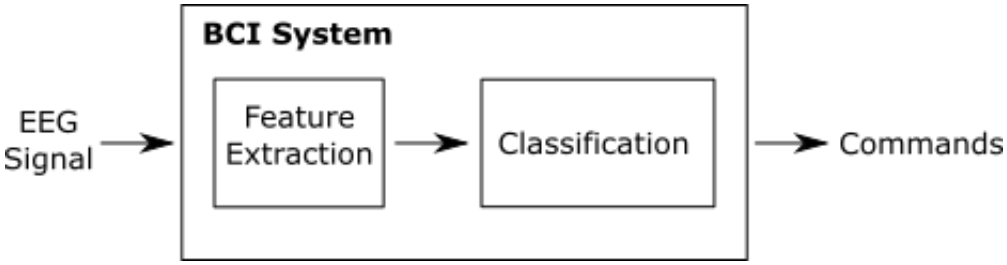


Figure 4.1 – BCI system general architecture. The BCI system is responsible for receiving an EEG signal and outputting a command, which should have strong correlation with the user's intent.

The general architecture of the two developed BCI systems is presented in Figure 4.1. It has two essential modules, the Feature Extraction module and the Classification module. The first module is responsible for extracting, using mathematical methods, a feature vector from an incoming EEG signal. The method used in the extraction is heavily dependent on the paradigm considered.

This feature vector is then used in the Classification module to obtain an identification of commands, which are independent from the type of paradigm considered. These two independent modules together serve to extract signal features that should have strong correlations with the user's intent, and to translate them into commands.

4.2. SSVEP-Based BCI system

In this section, we discuss the main modules of the SSVEP-Based BCI system, focusing on the methods used in the feature extraction module and in the classification module, as well as the pre-processing that is applied to the signal prior to these modules.

The methods described in this section were developed, initially tested and refined using the "AVI SSVEP Dataset"[11], a free dataset containing EEG measurements from healthy subjects being exposed to flickering targets in order to trigger SSVEP responses.

As described before, SSVEP needs an elicitation frequency for each command. Since we are using four commands, we therefore need four elicitation frequencies. These frequencies are then used to control four flickering targets on an LCD monitor, using the square wave modulation method.

[11] This Dataset was made by Vilic et al as part of his work on SSVEP-based BCI systems [55] and can be freely accessed (http://www.setzner.com/avi-ssvep-dataset/).

As stated before in Section 3.1.2 the lower frequencies (< 5Hz) are not practical to use in an SSVEP paradigm due to noise and high response time. Moreover, some combinations of frequencies should be avoided since the fundamental frequency can elicit stronger responses in the second and even third harmonic, producing confusion if the second harmonic of any given fundamental frequency is coincident with another fundamental frequency. Based on this, we chose the four fundamental frequencies represented in Figure 4.2, which are as follows:

$$H0 = \{10; 11; 12; 13\}\ \text{Hz} \quad \text{Eq. 4.1}$$

And the corresponding harmonics are:

$$H1 = \{20; 22; 24; 26\}\ \text{Hz} \quad \text{Eq. 4.2}$$

$$H2 = \{30; 33; 36; 39\}\ \text{Hz} \quad \text{Eq. 4.3}$$

The decision to use both the fundamental and the first harmonics arises from an analysis of the literature [2], [55]. Most of the methods studied consider both the fundamental and first harmonics as a mean to allow for the classification algorithm to be able to resolve between close related frequencies. The usage of more frequencies after the first harmonics becomes unnecessary, due to the exponential decay of the power of the frequency spectre for higher frequencies.

This set of frequencies determines the FFT resolution (spacing between the frequencies) to be a minimal of $\Delta R_{fft} = 1$ and the frequency resolution (minimum spacing between two frequencies that can be resolved) to be $\Delta R_f = 1$. The FFT resolution, together with the sampling frequency, influences the number of points used in the FFT algorithm, which can be modified by zero-padding the signal. The frequency resolution, on the other hand, directly influences the minimal length of the EEG signal, or Epoch, which in this case is the inverse of the frequency resolution, or $T = 1$s.

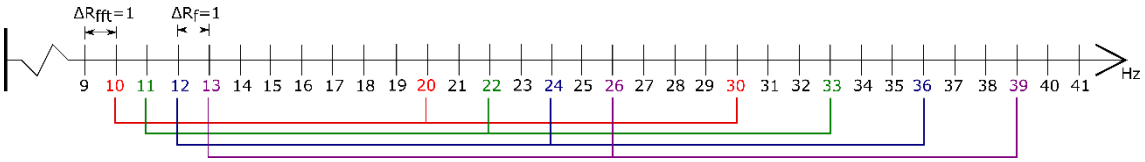


Figure 4.2 – Graphical representation of the fundamental and harmonic frequency sets, as well as the corresponding FFT resolution (ΔR_{fft}) and frequency resolution (ΔR_f). Each set of fundamental and harmonic frequency is represented in a different colour. The FFT resolution is given by the smallest frequency unit in the vector, which in this case is 1. The frequency resolution is given by the smallest spacing between adjacent frequencies, which in this case is 1.

In our work, we used a 4s length Epoch, which is consistent with those used in the literature, and an FFT resolution of 0.001Hz, by zero-padding the signal. This value of Epoch length provides a good trade-off between accuracy and response time, when compared to 1s length Epochs that provide high response time, but low accuracy and 6s Epochs that provide lower response time but higher accuracy. Regarding the FFT resolution, although this value is much lower than those used in the literature, our stimuli delivery system is robust enough to allow us to narrow the PSDA resolution peaks, allowing for better frequency separation.

4.2.1. Pre-processing

Upon arrival from the recording phase and prior to being sent to the BCI system, the raw signals from the occipital electrodes O1 and O2 were selected and the remaining information was discarded. This pair of electrodes was chosen in accordance with the works developed in the field of SSVEP-based BCI systems, more importantly the work of Pastor et al [36]. The signals were then detrended (i.e. their mean value or linear trend was removed) and filtered using a 15th order forward-backward Butterworth bandpass filter, between 8 and 50Hz. This method was applied to the signal in each electrode, separately, and then the signals were merged together in a single time vector, using "spatial average", i.e. the signals from both electrodes were averaged.

4.2.2. Feature Extraction

In our algorithm we used the Power Spectral Density-based Analysis (PSDA) method. This method is applied to an Epoch, using Eq. 4.4:

$$S = \frac{|X|^2}{N_{fft}} \qquad \text{Eq. 4.4}$$

This algorithm consists of multiplying the FFT of the Epoch with its complex conjugate ($X \cdot X^* = |X|^2$), and normalizing the result by dividing it by the number of data points after zero-padding (N_{fft}). In summary, it allows for easy extraction of the Power Spectral Density of each frequency, a single value that represents the contribution of each frequency in the considered signal Epoch, therefore also representing the "feature" for each particular frequency. These values are then used to identify which frequency has the highest contribution to the Epoch, from the list of command frequencies ($10, 11, 12, 13$ Hz).

4.2.3. Classification

We chose a classification method based on classes. The extracted features from a given Epoch (the PSDA values) are organized in four classes, each representing one of the stimuli frequencies. The absolute value attributed to each class, i.e. score, is calculated according to a modified version of Eq. 3.13 (see Eq. 4.5).

$$C_x = \sum_{H0-0.250}^{H0+0.250} |S| + \sum_{H1-0.250}^{H1+0.250} |S| + \sum_{H2-0.250}^{H2+0.250} |S| = C_{0x} + C_{1x} + C_{2x}; x = \{1, 2, 3, 4\} \quad \text{Eq. 4.5}$$

This method is similar to the one developed by Vilic et al [55], with the difference that ours uses the values from the PSDA, instead of the FFT values directly, has a higher interval of neighbouring frequencies usage, and considers the first two harmonics, instead of just the first harmonic. The usage of the neighbouring frequencies allows our algorithm to consider situations where the frequency contribution is scattered due to possible approximations. This is also a situation where the usage of the first and second harmonics is beneficial, because the fundamental frequencies are somewhat close together, but when we consider their harmonics, these are further apart and allow to better separate all the frequencies. On top of this, our stimuli system uses a square wave stimuli method. This is important to the considered harmonics due to the duty cycle of a square wave always being 50% or ½, in other words, every second frequency after the fundamental is no present. This lead us to include the second harmonic (third frequency from the fundamental) in our classification algorithm.

In summary, our method takes the values of the PSDA for each of the fundamental frequencies, adds the values for the frequencies, 0.250Hz immediately below and above the fundamental frequencies ($\text{H0} \pm 0.25$) and then does the same for each of the harmonics. The result is a single value that represents the class. For example, considering the class C_1, this class is composed by the sum of the values from the fundamental frequency 10Hz (C_{01}) and its harmonics 20Hz (C_{11}) and 30Hz (C_{21}). Thus, we sum the values of the PSDA between $9.750 - 10.250\text{Hz}$, $19.750 - 20.250\text{Hz}$ and $29.750 - 30.250\text{Hz}$, considering an FFT resolution of 0.001Hz (see Figure 4.3).

The values in all classes are then normalized in respect to each other, and then the dominant class is accepted, i.e. classified, has being the class chosen by the system.

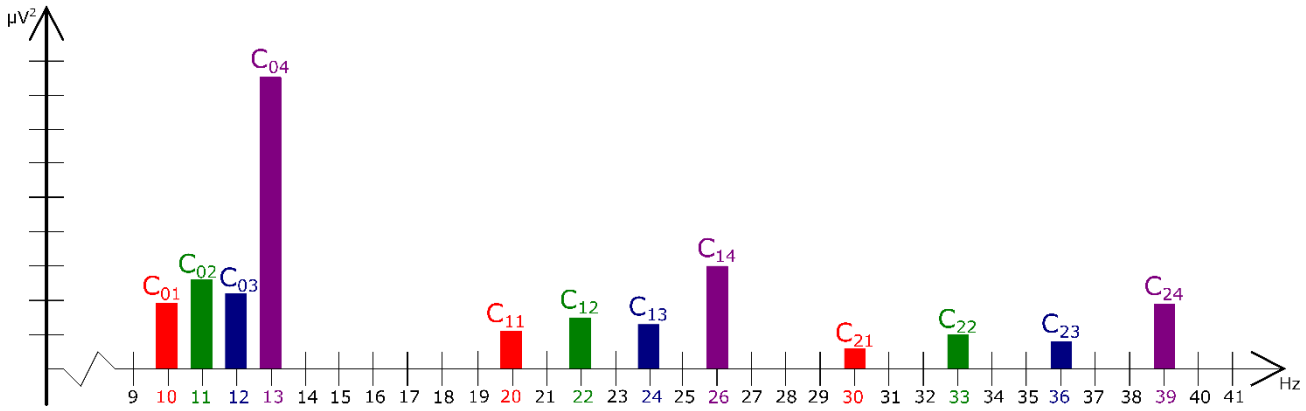


Figure 4.3 – Graphical representation of a PSDA example, highlighting the various class groups in different colours. In this example, we can clearly see that the dominating class would be the fourth, corresponding to the colour fuchsia ($13Hz$).

4.3. P300-Based BCI system

In this section, we discuss the main modules of the P300-Based BCI system, focusing on the methods used in the feature extraction module and in the classification module, as well as the pre-processing that is applied to the signal prior to these modules.

The methods described in this section were developed, initially tested and refined using the "Rodrigo Quian Quiroga EEG, ERP and single cell recordings database #4: Pattern Visual Evoked Potentials"[12], a free dataset containing EEG measurements from healthy subjects being exposed to target stimulation with an oddball paradigm to elicit a P300 response.

Considering the normal evolution of a P300 EEG signal (see Figure 4.4), we considered each Epoch to have a length of 600ms, beginning on the onset of the corresponding stimuli. Contrary to the SSVEP paradigm, where each Epoch can be used to analyse and classify the signal into each of the 4 considered commands, the P300 paradigm requires each command to be associated with its own Epoch. This characteristic arises from the fact that a P300 eliciting stimuli is time locked to the EEG signal. Considering the oddball stimuli system, and the fact that we are using four commands, this means that our systems is basically analysing each Epoch that is time locked to one of the four stimulus, and decides which one of them is the oddball, i.e. the P300 EEG signal.

[12] This Dataset was made by Quiroga et al as part of his work on P300-based BCI systems [74] and can be freely accessed (https://vis.caltech.edu/~rodri/data.htm).

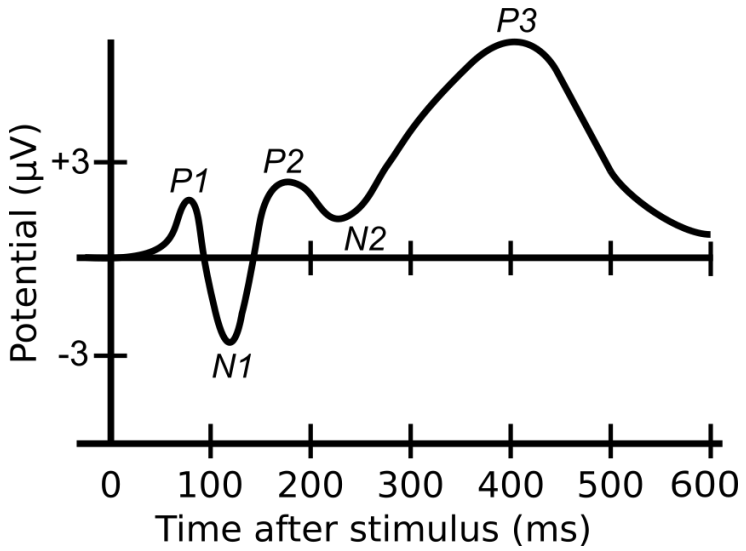


Figure 4.4 – Graphical representation of the temporal evolution of a P300 EEG signal. Notice the pronounced negative deflection at $\sim 100ms$ (N1 or N100), followed by the characteristic positive deflection at $\sim 300ms$ (P3 or P300).

4.3.1. Pre-processing

As it happened with the SSVEP signals, upon arrival from the recording phase and prior to being sent to the BCI system, the raw signals from the electrodes $F3, FC5, FC6$ and $F4$ were selected, taking into account the results obtained from the technical report presented by H. Ekanayake [73], with regards to the usage of the Emotiv headset to record P300 signals. In this case, the remaining information from the other electrodes was used to improve the signal quality.

The signals from all electrodes were detrended (i.e. their mean value or linear trend was removed) and filtered using a 4th order forward-backward Butterworth bandpass filter, between 1 and 40Hz. Then the signals were windsorized (i.e their value below and above two percentiles were replaced by the values on those percentiles) below the 10th percentile and above the 90th percentile, in order to minimize the effect of eye blinks, eye movement, muscle activity, or user movement artifact outliers in the signal [75]. A spatial filter was built using all the electrodes to reduce the contribution of all the electrodes in a specific electrode, better isolating its information. Considering the available electrodes, our chosen filter was the Common Average Referencing (CAR) [1].

Spatial filters derive signal features by combining data from two or more locations so as to focus on activity with a particular spatial distribution. CAR spatial filters apply a fixed set of weights to a linear combination of channels (i.e. electrode locations) in order to improve the signal-to-noise ratio (SNR) of EEG signals [1].

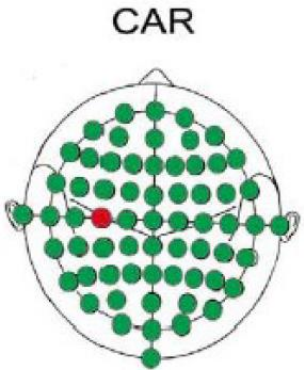


Figure 4.5 – Visual representation of the Common Average Referencing (CAR) spatial filter. In this example we see the output electrode in red and all the remaining electrodes in green. All the electrodes are averaged and the mean is subtracted from the output channel of interest. Adapted from [1].

In CAR (Figure 4.5), the mean value of the entire electrode montage used in the recording phase is subtracted from the signal of the channel of interest:

$$CAR\big(x_i(t)\big) = x_i(t) - \frac{1}{m}\sum_{i=j}^{m} x_j(t) \qquad \text{Eq. 4.6}$$

Where $x_i(t)$ is the time-dependant signal acquired by the i-th electrode and m is the total number of electrodes used in the recording.

This method was applied separately to the signal in each of the four interest electrodes ($F3, FC5, FC6$ and $F4$) and then the signals were merged together in a single time vector, using "spatial average", i.e. the signals from the four electrodes were averaged, and this single time vector was lastly scaled to the interval $[-1, 1]$.

4.3.2. Feature Extraction

As mentioned in Section 3.2.1, due to the inherent characteristics of a P300 response the methods used to extract and classify features are mostly centred on time domain processing. Given that the time domain of an EEG signal is more prone to artifacts and has a lower SNR, usually most researchers tend to combine multiple methods of feature extraction and classification, for better results. This, coupled with our intention to use Single Trial Recognition lead us to use a combination of two feature extraction methods: Area Analysis and Peak Picking.

Contrary to common practice in P300 BCI systems, we decided to use both feature extraction methods together as a voting system. This innovation improves accuracy and does not affect ISI or ITI, therefore it does not interfere with the system response time.

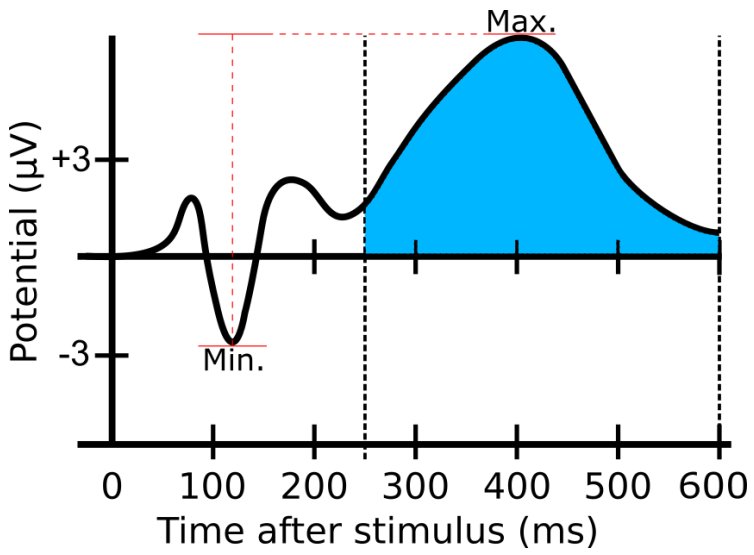


Figure 4.6 – Graphical representation of the temporal evolution of a P300 EEG signal, highlighting in light blue the zone defined for the Area analysis and in red the peaks used in the Peak Picking method. As mentioned in the text, the Area analysis method uses a defined temporal interval, between 250 and $600ms$, returning the total area under the curve for that interval. The Peak Picking method, on the other hand, searches the curve for its lowest value in the interval between 0 and $250ms$, and for its highest value between 250 and $600ms$, returning the difference in amplitude between these two peaks.

The Area analysis method consists of calculating the 'area under the curve' of the P300 component of the signal for each Epoch. We compute it for the interval between 250ms and 600ms, returning a single value or "feature" for each respective Epoch (see Figure 4.6).

The Peak Picking method, on the other hand, calculates the difference between the highest point of the signal between 250ms and 600ms and the lowest point of the signal between 0ms and 250ms. This returns a single value or "feature" for each respective Epoch (see Figure 4.6).

4.3.3. Classification

A low Inter Stimuli Interval (ISI) value allows for faster response times and high accuracy, in a Single Trial Recognition scenario. However, if the ISI value is too low, it can negatively affect the accuracy and generate false positives in the classification method, due to over superposition of the signals/stimuli. Using the mentioned dataset, and considering the ISI limitations, we developed a stimuli system that uses an ISI of 250ms, which gives us an Inter Trial Interval (ITI) of 1350ms or 1.35s (see Figure 4.7), lower than the 4s trials considered on the SSVEP system.

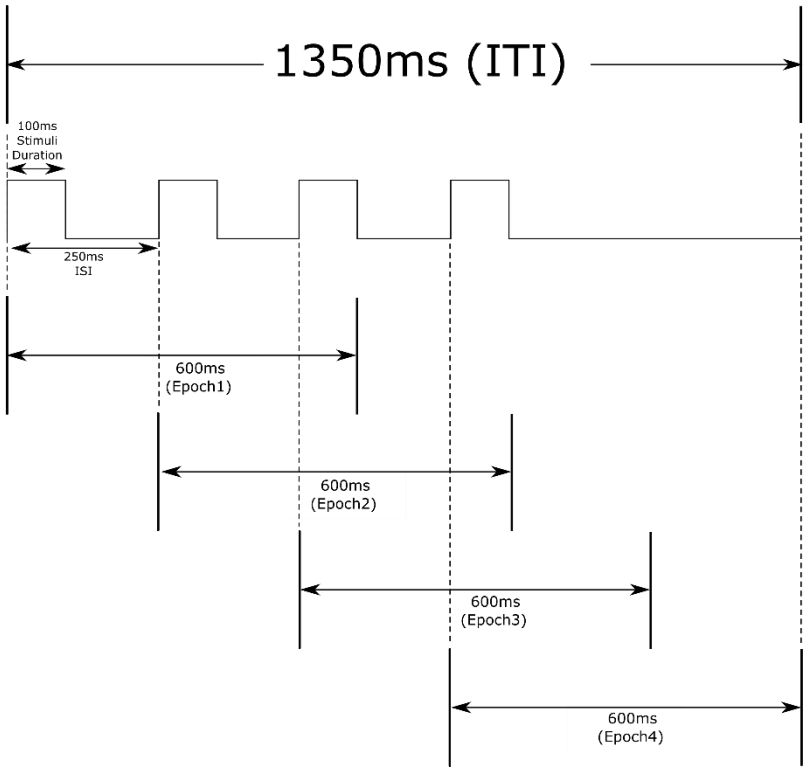


Figure 4.7 – Graphical representation of our P300 stimuli system. ITI – Inter-Trial Interval; ISI – Inter-Stimuli Interval.

The chosen classification method is a method based on classes, to some extent similar to the method used on the SSVEP system. The results from each of the two methods on the feature extraction system generates four normalized feature vectors per trial, on a total of ten trials per system's decision (i.e. each system's decision takes ten 1350ms trials, or 13,50s of EEG signal, for a total of 40 epochs of 600ms, considering the stimuli system in Figure 4.7). Each feature vector corresponds to one of the "directional" Epochs, and these get sorted in such a way that when the group of 40 epochs is put together, there is a group of 10 epochs correspondent to the "Left" movement, 10 to the "Up" movement, 10 to the "Right" movement, and 10 to the "Down" movement (see Figure 4.8**a**).

The feature vectors are then summed together and normalized, in a process termed weighted voting. This process considers the weight of each of the four features per trial, that have a score between 0 and 1. When the ten trials are summed, inside each epoch group, it is expected that the group of epochs that has the highest value is the one correspondent to the expected movement (this happens when the feature extraction algorithm labels the same movement epoch with the highest, or close to highest, score in most of the ten considered trials). These summed values are finally normalized and represent the weighted vote from each of the two feature extraction systems (see Figure 4.8**b**).

The weighted vote vectors are finally combined together through the same process, this time considering the weight of each feature extraction system, and normalized, creating the final decision vector (see Figure 4.8**c**) and leading the system to output the highest valued epoch (i.e. the one with the score of 1) as the final decision.

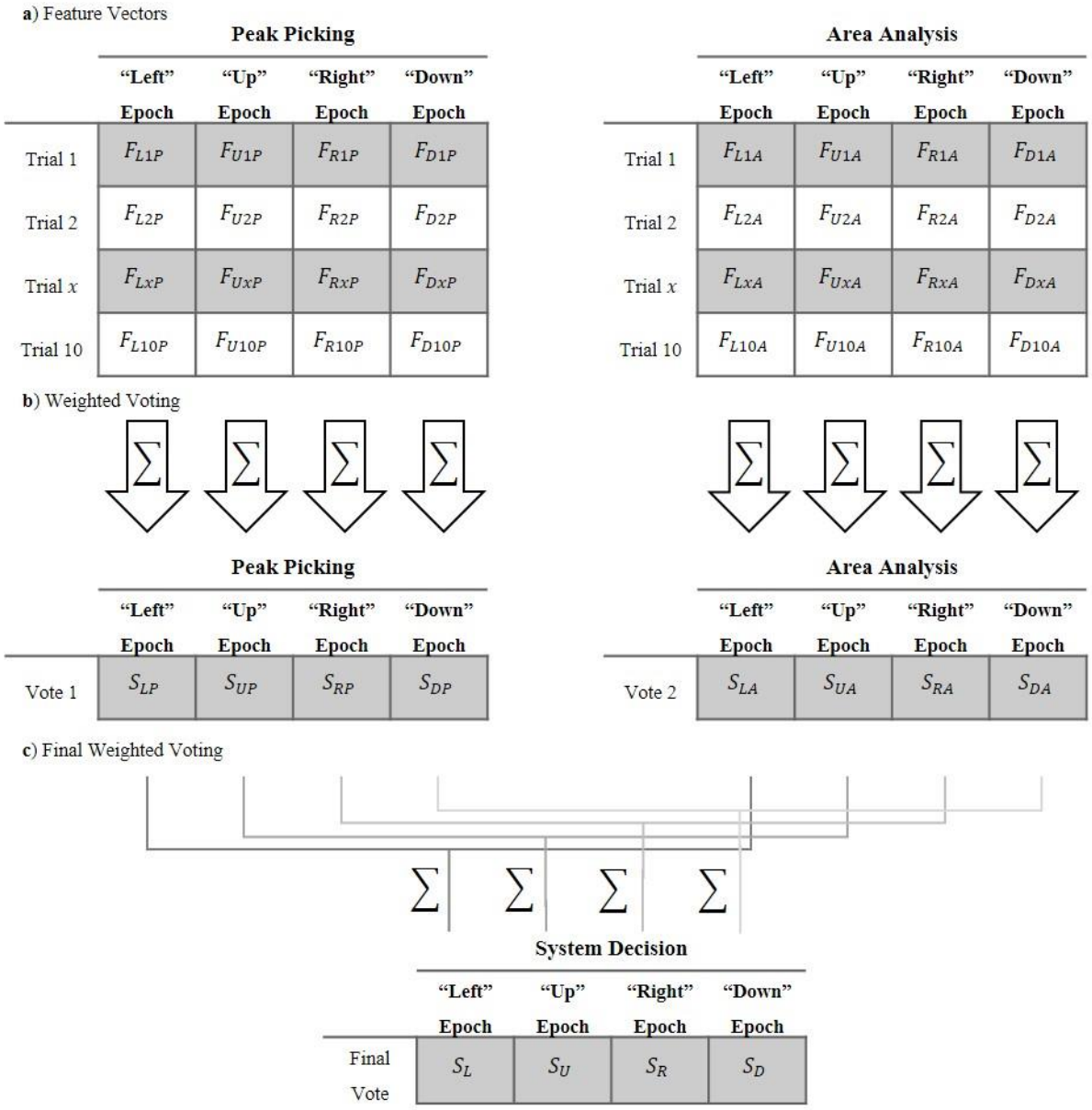


Figure 4.8 – Schematic representation of the Classification algorithm for the P300-based BCI system. **a**) Feature vectors from both systems, divided by trials and movement epochs, according to the stimuli system; **b**) Weighted voting system, that takes the scores from the ten trials and sums them, per considered movement epoch; **c**) Final Weighted Voting system, that takes the scores from the two distinct feature extraction algorithms and sums them, per considered movement epoch, outputting a final voting score. The highest score gets chosen as the system decision; F – Feature Score; L – Left; U – Up; R – Right; D – Down; P – Peak Picking; A – Area Analysis; S – Score. It is worth noting that all values presented in the tables (F and S) are always normalized per table line

4.4. Summary

In this section we discussed the methods used on our BCI systems, focusing on the pre-processing methods, and on the feature extraction and classification modules built for both the SSVEP and P300 paradigms.

The SSVEP system used a PSDA method to extract features from the signal and a classification based on classes to return a decision based on the characteristics of the signal. The P300 system on the other hand used two methods to extract features from the signal, Area Analysis and Peak Picking, and also used a classification based on classes, that combined the information from the feature extraction phase that was applied to each trial and each method, and returned a decision based on the characteristics of the signal.

5. BCI Application

In this chapter we present and discuss the developed application, that uses the already described BCI systems. This application is based on virtual environment navigation, more precisely on two dimensional maze navigation, where the user needs to make decisions on the directions that will eventually take the avatar to the goal (i.e. the exit of the maze).

In the following subsections, we present and discuss the general architecture of the interface, the distinct stimuli inherent to the two different BCI systems, and the movement of the avatar. We then present the types of mazes developed, how the user interacts with them, and finally the technology used to develop the final system.

5.1. Architecture

The general architecture of the developed application is presented in Figure 5.1. It is composed by the already described BCI system that receives an EEG signal from the user and outputs a command. This command is then used by the application ("User Interface") to move the avatar inside the maze. When a decision is necessary, the application outputs specific stimuli, which are perceived by the user, eliciting distinct EEG signals that are used in the BCI system to output more commands, closing the cycle.

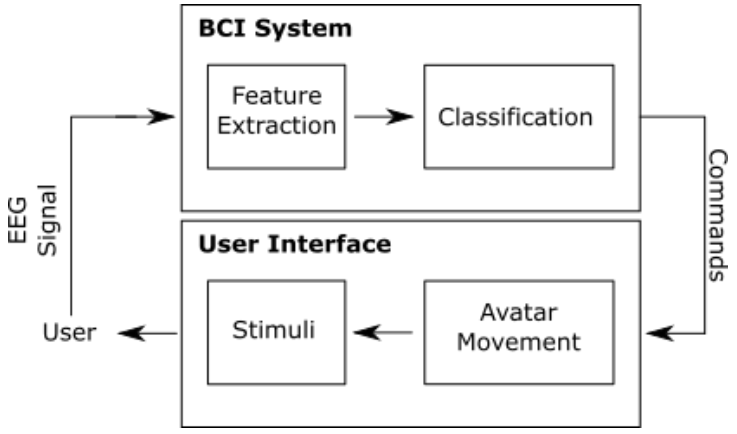


Figure 5.1 – General architecture of the interface application. As mentioned in the previous chapter, the BCI system is responsible for receiving an EEG signal and outputting a command, which should have strong correlation with the user's intent. The User Interface then uses that command to move the avatar and then outputs stimuli that are perceived by the user and elicit distinct EEG signals, which are then used by the BCI system, closing the system.

Depending on the type of paradigm considered, the outputted stimuli will be different and will affect the generated EEG signal. Due to this fact, we essentially developed two distinct applications, one for the SSVEP system and another for the P300 system. It is worth nothing however, that the avatar movement and the mazes themselves are common to both interfaces.

5.2. Stimuli

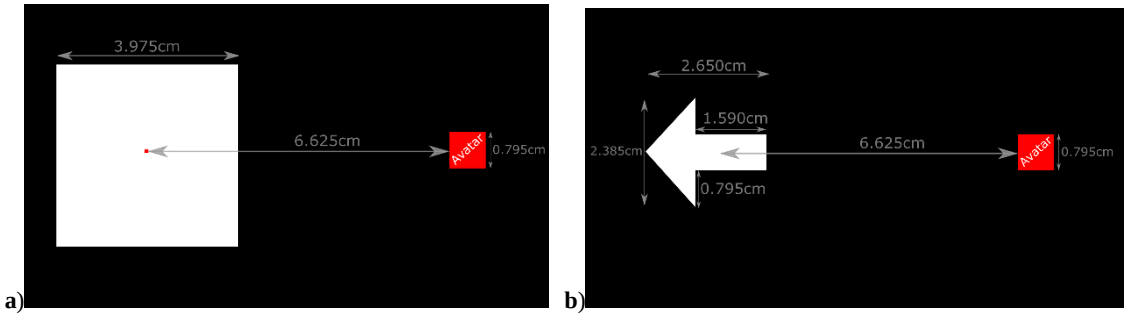


Figure 5.2 – A detailed scheme of the stimuli system for both the SSVEP and P300 systems, depicting the general distances and dimensions of the targets. **a**) SSVEP system; **b**) P300 system. It is worth noting that these are just simplified representations of the system, i.e. only one of the four total targets is represented.

The stimuli consist of four targets around the avatar, as shown in Figure 5.2. It is worth noting that although the distances between the avatar are similar for both the SSVEP and P300 system, the way in which the stimuli are presented, and the stimuli itself varies between these two systems.

These targets are used to elicit an EEG signal response on the user. To do this, whenever the system needs a decision to be taken by the user, it activates the stimuli system, which presents the four targets. The BCI system then analyses the elicited EEG signal and outputs a command, leading to the movement of the avatar.

5.2.1. SSVEP Stimuli

The SSVEP stimuli is elicited using four flickering squares, with each square having a total area of $\sim 15.75 cm^2$, more than enough to cover the fovea, and the distance between any adjacent targets being at least $\sim 9.34 cm$ (calculated as the shortest diagonal distance). These dimensions were designed so that at any given point, the fovea can only cover one square (see Figure 5.2**a**).

An SSVEP signal is usually elicited following excitation of the retina by a visual stimulus ranging from 3.5Hz to 75Hz. The brain then generates electrical activity at the same (or multiples of the) frequency of the visual stimulus [40].

As mentioned in Section 4.2, we chose four fundamental frequencies (10Hz, 11Hz, 12Hz and 13Hz) for the SSVEP BCI system. These frequencies were used in the application stimuli system, to control the flickering of the square targets, so that each directional square would flicker at a distinct frequency from each other (see Figure 5.3).

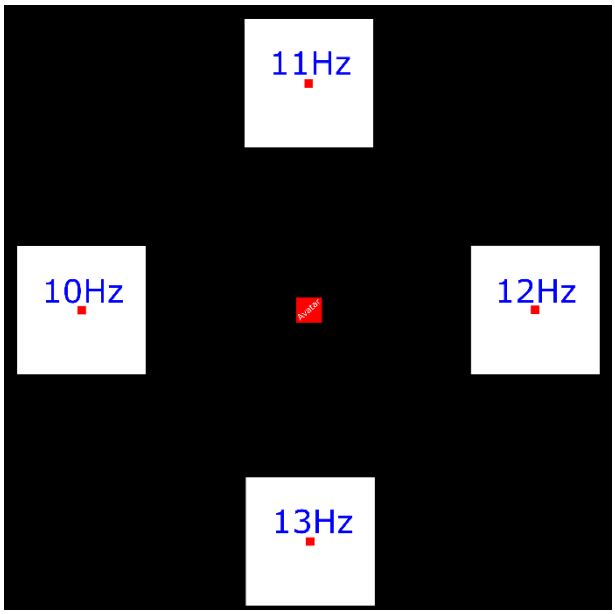


Figure 5.3 – Representation of the SSVEP stimuli system, depicting the frequency of each target and the general position around the central avatar. The smaller red squares in the centre of the targets are used to ease the user focus, and are called gazing points in the purpose of this study.

The method used in the flickering of the targets was a modified version of Eq. 3.19, adapted to include an alpha modulation value:

$$Alpha(f,i) = 255 \times \frac{\left(1 + square\left(2\pi f \frac{i}{R}\right)\right)}{2} \qquad \text{Eq. 5.1}$$

The alpha value is part of the Red Green Blue Alpha (RGBA) colour space model, used computer graphics representation. The alpha channel in particular is used as an opacity channel for pixels, i.e. if a pixel as an alpha value of 0 it is fully transparent, whereas a value of 255 gives a fully opaque pixel. By using a white square on a black background this method can be used to modulate the opacity of the square between 0 and 255, making it flicker with the intended frequency.

5.2.2. P300 Stimuli

A P300 signal is usually elicited with a two-stimulus discrimination task. This procedure has been dubbed the 'oddball paradigm', since two stimuli are presented in a random series such that one of them occurs relatively infrequently, i.e. the oddball [43].

In the case of this application, we present each stimulus arrow in a randomized order, such that the "target" stimulus would always be the oddball. This presentation is done as mentioned in Section 4.3.3, with each stimuli being presented for 100ms followed by a 150ms interval (see Figure 5.4), and with the presentation sequence being randomized in each trial. To improve the oddball paradigm, each stimulus is represented as a distinctively coloured arrow (White, Red, Green and Blue).

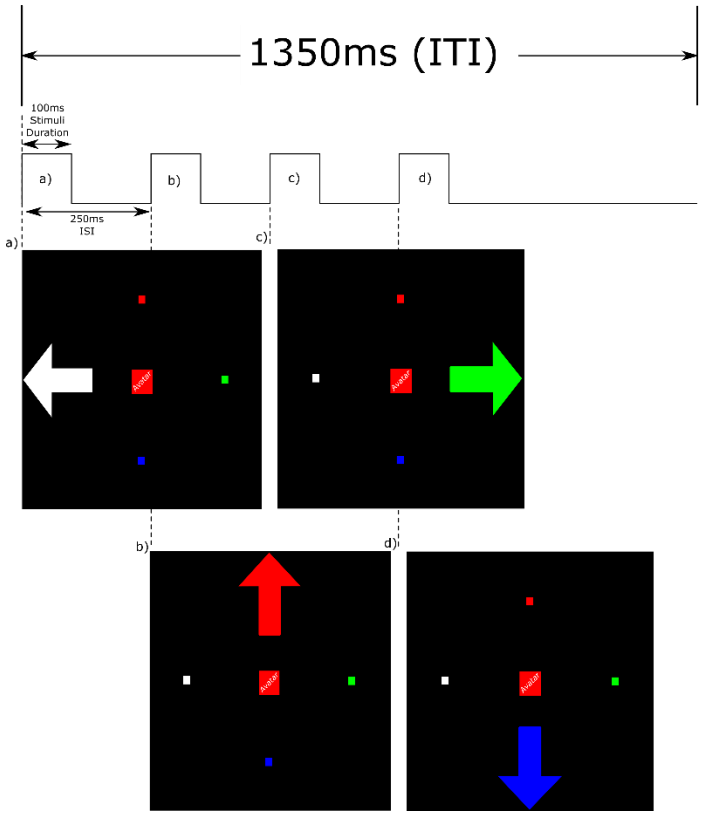


Figure 5.4 – Schematic representation of the P300 stimuli system timeline. Each arrow is presented during the stimuli duration period, in a randomized sequence, followed by a pause period where no stimuli is present. The smaller squares, which colours matches the correspondent arrow, are used to ease the user focus, and are called gazing points in the purpose of this study.

5.3. Moving the Avatar

The movement of the avatar is done in a straight trajectory, between "decision points" (DP). The avatar typically starts on the entrance of any given maze and activates the stimuli system to request for a command from the BCI system. The BCI system then returns a command that is translated into a movement in any of the four directions, represented by the four stimuli arrows. The avatar initiates the movement, during which the user rests and decides on the next direction to take. When the avatar reaches another DP, and this DP is not the required goal (i.e. maze exit), the stimuli system is again activated and the cycle repeats, until the user guides the avatar to the goal.

The number of DPs and the available directions in a maze depend on the type/complexity of the maze itself. It is worth noting that if for any reason (user error or system error) the command returned by the BCI system correspond to an invalid direction (e.g. a maze wall) the command is registered in the log files, but the avatar does not initiate movement to a valid DP, instead it requests another command, by activating the stimuli system again.

5.4. Mazes

A maze is a path or collection of paths, typically from an entrance to a goal. The word is used to refer both to branching tour puzzles through which the solver must find a route, and to simpler non-branching ("unicursal") patterns that lead unambiguously through a convoluted layout to a goal [76]. In the case of our application, the used mazes are from the first type, a branching puzzle where the user must find a route from the entrance to a goal.

Since the main purpose of this application is to provide a way to test and evaluate the developed BCI systems, the types of mazes used heavily depend on what is being measured/evaluated.

5.4.1. T-shaped Maze

The T-shaped maze (Figure 5.5) is a simple maze, shaped as a plus "+" symbol. This maze allows the user to perform single "T" decisions, i.e. the user will always enter the centre of the maze from a direction which will not contain the goal, and therefore needs to choose from the other three directions. This type of maze allows us to gradually introduce the user to the user interface (UI) and to the movement mechanics, while also allowing for an initial measurement of user's accuracy.

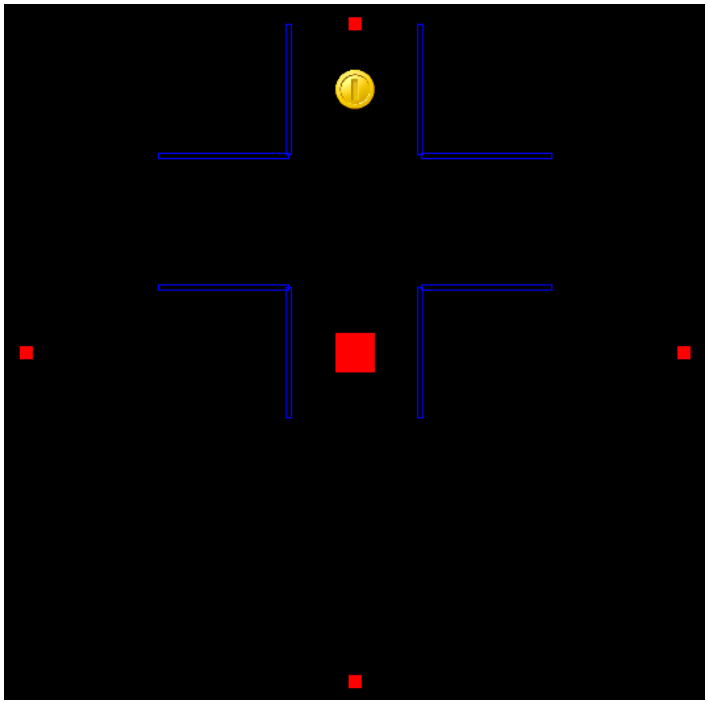


Figure 5.5 – A screenshot image of the T-shaped maze used in the BCI systems evaluation. The goal of this maze is to move the red squared avatar towards the yellow circular-shaped coin. This is done by moving the avatar up towards the middle DP and then again up, towards the goal. When the avatar reaches the goal, the goal changes place and then the user must reach it again, in a total of 16 movements of the avatar. It is worth noting that, although in this image the interface is the one seen in the SSVEP system, the maze itself can be used in both systems.

5.4.2. Full Maze

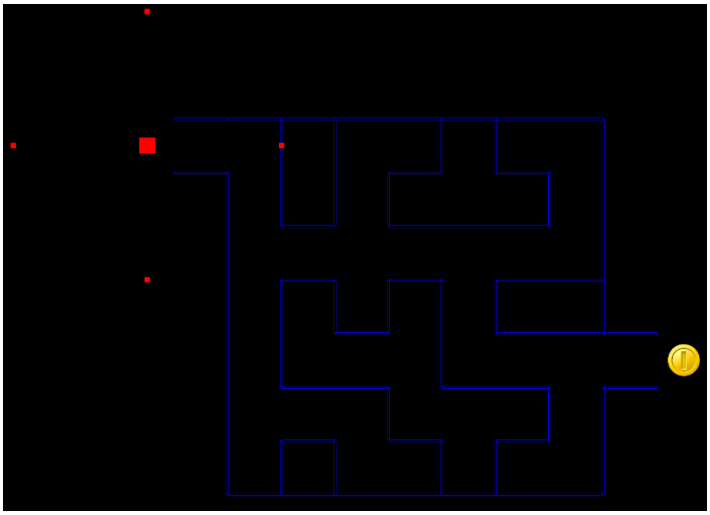


Figure 5.6 – A screenshot image of the Full maze used in the BCI systems evaluation. The goal of this maze is again to move the red squared avatar towards the yellow circular-shaped coin. This is done by moving the avatar through the maze, choosing the correct path, towards the goal. When the avatar reaches the goal, the goal changes place to the previous entrance of the maze, and then the user must reach it again. It is worth noting that, although in this image the interface is the one seen in the SSVEP system, the maze itself can be used in both systems.

The Full maze (Figure 5.6) is a full sized maze, with 13 decision points, one entrance and a exit/goal. This maze allows the user to perform multiple "+" decisions, i.e. the user will have the freedom to choose any direction, from the four available, at each decision point throughout the maze. This type of maze introduces the user to the usage of the BCI system in a more complex scenario, while also allowing for a measurement of user's accuracy in a distinct scenario.

5.5. Technology

The key technology, and corresponding versions, used in the making of this work were: MATLAB R2014a x86; Psychophysics Toolbox 3.0.11; and the Emotiv Headset Application Programming Interface (API) together with Microsoft Visual Studio Professional 2013 update 5.

The BCI system and user interface were both developed in MATLAB, a high-level language and interactive environment for numerical computation, visualization, and programming. The language, tools, and built-in math functions of this environment enables the exploration of multiple approaches and reaching a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java. MATLAB can be used for a range of applications, including signal processing and communications, image and video processing, control systems, test and measurement, computational finance, and computational biology.

Furthermore, the user interface was built with a toolbox specially designed for vision and neuroscience research, the Psychophysics Toolbox Version 3 (PTB-3). This free set of functions interfaces between MATLAB and the computer hardware, providing access to the display frame buffer and colour lookup table, allowing synchronization with the vertical retrace, supporting sub-millisecond timing, exposing raw OpenGL commands, supporting video playback and capture as well as low-latency audio or hardware triggers, and facilitating the collection of observer responses through regular and specialized input devices, among other functionalities [77], [78].

Finally, the communication between MATLAB and the Emotiv headset was done using the Emotiv System Developer Kit (SDK), which comes with a set of built-in library functions, or API, that includes routines, data structures, object classes and variables that serve as tools to build software applications between the Emotiv and the chosen programming environment.

Since our intention was to integrate all the different modules into a single system, there were some obstacles regarding the chosen programming environment. The usual documented method for extracting raw signals from the Emotiv headset involves a constant communication with the headset, therefore storing each group of data packets as they arrive at the computer, as described in the following pseudocode example:

```
Connect to Emotiv
Check if the connection was successful
Create a memory buffer
Set the memory buffer size to fit the packet size
while(true)
          if Check for Errors equals "All Ok"
               Get Event type
               Get User ID
               if Event type is User Added
                    if User ID not null
                         Enable Data Acquisition
          else if Error
               break
          if Data Acquisition is Enabled
               Update the Emotiv Handle
               Get the number of samples recorded
               if Number of samples not null and more than zero
                    Store each sample from each electrode
end
Disconnect from Emotiv
Free the memory buffer
```

Although this method works and is essential in experiments where there is the need to collect variable lengths of EEG data within the same trial, the constant communication requires the existence of a loop in MATLAB that conflicts with the fact that while the communication is occurring, the stimuli system also needs a loop that is responsible for changing the stimuli that is constantly being drawn to the graphical buffer of the computer. The need for two loops could have been overcome using the MATLAB functionality that creates multiple threads or "workers" that run in parallel, called the Parallel Computing Toolbox. The problem is that these "workers" do not have access to the Graphical Processing Unit (GPU) memory buffer, and therefore the PST-3 stimuli system cannot run from inside a worker. Our solution was to create a memory buffer big enough to store the whole intended length of the EEG data and therefore removing the need to constantly update and store the arriving samples, as described in the pseudocode below. This method works in the case of our work due to the fact that our trials have fixed lengths, so we know *a priori* the size of the memory buffer.

```
Connect to Emotiv
Check if the connection was successful
Create a memory buffer
Set the memory buffer size to fit the trial acquisition time
while Event type is not User Added
     Get Event type
     Get User ID
     Enable Data Acquisition
end
while Event type is not State Updated
     Get Event type
end
Update the Emotiv Handle
Start the PST-3 Stimuli System
Update the Emotiv Handle
Get the number of samples recorded
Store each sample from each electrode
Disconnect from Emotiv
Free the memory buffer
```

The stimuli were presented on a 27-inch Dell UltraSharp 2709W widescreen flap panel LCD monitor with a resolution of 1920×1200 and a refresh rate of $60Hz$. This monitor was connected via a Video Graphics Array (VGA) cable to an Asus ROG GL552J laptop, with a i7-4720HQ CPU @$2.60GHz$ and $8.00Gb$ of RAM, running a Windows 8.1 x64 OS.

5.6. Summary

In this section we discussed the developed user interface used to evaluate the BCI systems, focusing on: the general architecture of both systems; how the stimuli was delivered to the user and the main differences in the SSVEP and P300 systems; the movement of the avatar inside the interface; the two distinct mazes, their differences and what was the purpose of them; and finally the technology that allowed everything to work together, as well as some of the difficulties that come with the intent of making a single unified system that incorporates everything.

6. Experimental Evaluation

In this chapter we present and discuss the two phases of the experimental evaluation of the BCI application, as well as the main results.

The purpose of this evaluation was to compare the performance of our BCI systems for a "plug and play" approach using high quality datasets, collected with non-portable EEG recording systems in a lab environment, and using data from users, collected with a commercially affordable portable recording device (i.e. Emotiv) in an out-lab environment.

We conclude the chapter by comparing and discussing the main results achieved in the two experiments.

6.1. Evaluation using High Quality Datasets

Prior to using the developed BCI systems with the data recorded from the Emotiv headset, these were tested using freely available high quality datasets. It is worth noting that these tests were used more as a "proof of concept" than as a mean of evaluating our BCI systems, mostly due to the different recording conditions and stimuli delivery that were used to collect the data.

6.1.1. SSVEP Dataset

The SSVEP dataset used for testing the SSVEP BCI system was a dataset made by Vilic et al [55] as part of his work on SSVEP-based BCI systems, and can be freely accessed from his website[13].

The EEG signals from the dataset were recorded from seven healthy subjects, four males and three females age 28.1 ± 2.3 years, using three gold plated electrodes, placed along the subject's scalp: a ground electrode at Fpz; a reference electrode at Fz; and a signal electrode at Oz. The impedances were kept at $5\text{k}\Omega$ or lower, and the signal was amplified using a g.USBamp from g.tec set to a sampling rate of 512Hz and filtered using an analog notch filter at 50Hz.

The stimulus was delivered using an LCD monitor with a refresh rate of 120Hz, and the experiment was split into two phases: the single target stimulation and the multi target stimulation.

[13] http://www.setzner.com/avi-ssvep-dataset/

Single-Target Stimulation

During the single-target stimulation the users, a total of four from the pool of seven, were sit at $\sim 60\text{cm}$ from the monitor and asked to gaze at a single flickering target, changing color from black to white, with a side dimension of 1.7cm (area of 2.89cm^2) for 30s, after which the subjects would take a brief break. The flickering frequencies used were $\{6, 6.5, 7, 7.5, 8.2, 9.3, 10\}$ Hz, and the 30s trials were repeated at least three times for each frequency. Given that our SSVEP BCI system is built to accept 4s epochs, we took the 30s trials and split them into 4s epochs, discarding the last 2s of data, which accounts for a total of seven epochs per trial. So considering everything, we ended up with a total of 665 epochs of 4s, across all subjects, which accounts for: 126 epochs corresponding to a target frequency of 6Hz; 91 epochs corresponding to 6.5Hz; 105 epochs corresponding to 7Hz; 84 epochs corresponding to 7.5Hz; 91 epochs corresponding to 8.2Hz; 84 epochs corresponding to 9.3Hz; and 84 epochs corresponding to 10Hz. It is also worth noting that the total amount of epochs is not uniform across subjects due to the fact that each subject had a different number of data trials.

The SSVEP BCI system algorithm parameters were adjusted to the dataset, maintaining the intended values of epoch length (i.e. 4s), FFT resolution (i.e. 0.001Hz), and the same feature extraction (i.e. PSDA) and classification methods (i.e. Classes). The only parameters that differed from the ones in our user evaluation were the sampling rate (512Hz vs 128Hz), the number of classes (7 classes vs 4 classes) and the refresh rate of the monitor (120Hz vs 60Hz).

In the following tables, we present the accuracy results of our algorithm considering each subject in particular (Table 6.1) and all the target frequencies across all subjects in general (Table 6.2).

Table 6.1 – Accuracy results from the intra-subject analysis of the single-target dataset, with the developed SSVEP BCI system algorithm. In the context of this work, accuracy per subject is defined as the number of correct decisions over the total number of decisions and the overall accuracy is a direct mean of the latter.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	189	113	59.79
2	182	124	68.13
3	147	96	65.31
4	147	105	71.43
$Mean \pm std$	166.3 ± 22.41	109.05 ± 11.90	66.17 ± 4.931

Table 6.2 – Accuracy results from the inter-frequency analysis across all subjects and trials of the single-target dataset, with the developed SSVEP BCI system algorithm.

Freq. (Hz)	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
6	126	95	75.40
6.5	91	67	73.63
7	105	64	60.95
7.5	84	71	84.52
8.2	91	11	12.09
9.3	84	68	80.95
10	84	62	73.81
$Mean \pm std$	95.00 ± 15.58	62.57 ± 25.28	65.91 ± 24.85
$Mean \pm std$ $w/o\ 8.2Hz$	95.67 ± 16.95	71.17 ± 12.09	74.88 ± 8.085

From an analysis of Table 6.1, we can see that users reported a mean value of accuracy of 66.17%, with a minimal value of 59.79% on user #1. From an analysis of Table 6.2, we can also see that the accuracies across all used frequencies were similar (around $70\% - 80\%$), except for the frequency of 8.2Hz, which achieved the lowest accuracy of 12.09%. Since this frequency deviates from the others, we decided to also include an analysis of the mean accuracy without this frequency, and the system achieved a mean accuracy of 74.88% across the remaining frequencies. We suspect that the low accuracy value attributed to this frequency may be tied to the high FFT resolution used, which can affect frequency components, especially in situations where the stimuli system is not robust, i.e. the stimulus frequency is not accurately being delivered.

Multi-Target Stimulation

During the multi-target stimulation, the users, a total of five from the pool of seven, were sit at ~ 60cm from the monitor and asked to gaze at a flickering target, from an UI with seven simultaneous flickering targets, changing colour from black to white, all with a side dimension of 1.7cm (area of 2.89cm^2) and separated by a 2.1cm gap. Users had to look for 16s, after which they took a brief break and then were asked to change gaze to another target. The flickering frequencies used were $\{6, 6.5, 7, 7.5, 8.2, 9.3, 10\}$ Hz, and the 16s trials were repeated ten times for each frequency and each subject performed two sessions. As it happened with the single target dataset, we split the 30s trials into 4s epochs, which accounts for a total of four epochs per trial. We ended up with a total of 400 epochs of 4s, across all subjects, which accounts for: 160 epochs corresponding to a target frequency of 6Hz; 20 epochs corresponding to 6.5Hz; 40 epochs corresponding to 7Hz; 60 epochs corresponding to 7.5Hz; 100 epochs corresponding to 8.2Hz; and 20 epochs corresponding to 9.3Hz. Although the frequency of 10Hz was used to flicker one of the targets, there is no data in the dataset regarding the use of this frequency to evoke an SSVEP response. In spite of this, we still consider the frequency of 10Hz as one of the seven possible classes, maintaining the algorithm parameters discussed above.

In the following tables, we present a similar analysis to the single-target dataset, i.e. the accuracy results of our algorithm considering each subject in particular (Table 6.3) and all the target frequencies across all subjects in general (Table 6.4). It is worth noting that the subject IDs are arbitrary, that is, the same subject ID across both datasets (the single and multi-target) may not correspond to the same subject.

Table 6.3 – Accuracy results from the intra-subject analysis of the multi-target dataset, with the developed SSVEP BCI system algorithm.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	80	46	57.50
2	80	63	78.75
3	80	24	30.00
4	80	67	83.75
5	80	62	77.50
$Mean \pm std$	80	52.40 ± 17.78	65.50 ± 22.23
$Mean \pm std$ *w/o #3*	80	59.50 ± 9.256	74.38 ± 11.57

Table 6.4 – Accuracy results from the inter-frequency analysis across all subjects and trials of the multi-target dataset, with the developed SSVEP BCI system algorithm.

Freq. (Hz)	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
6	160	130	81.25
6.5	20	17	85.00
7	40	30	75.00
7.5	60	45	75.00
8.2	100	26	26.00
9.3	20	14	70.00
$Mean \pm std$	66.67 ± 54.65	43.67 ± 43.69	68.71 ± 21.58
$Mean \pm std$ *w/o* $8.2Hz$	60.00 ± 58.31	47.20 ± 47.88	77.25 ± 5.890

From an analysis of Table 6.3, we can see that users reported a mean value of accuracy of 65.50%, with a minimal value of 30.00% on user #3. Since this user deviates from the others, we decided to also include an analysis of the mean accuracy without this user, and the system achieved a mean accuracy of 74.38% across the remaining users. Regarding this user in particular, there is not enough information that can safely explain the low value of accuracy achieved. Most likely this value is due to low attention, on the user part.

From the analysis of Table 6.4, we can see that the accuracies across all used frequencies were similar (around 75%), except for the frequency of 8.2Hz, which achieved the lowest accuracy of 26.00%. As it happened with the last dataset, since this frequency deviates from the others, we decided to also include an analysis of the mean accuracy without this frequency, and the system achieved a mean accuracy of 77.25% across the remaining frequencies. Again, we suspect that the low accuracy value attributed to this frequency may be a combination of the high FFT resolution and a poorly designed stimuli delivery system.

6.1.2. P300 Dataset

The P300 dataset used for testing the P300 BCI system was a dataset made by Quiroga et al [74], [79] as part of his work on P300-based BCI systems, and can be freely accessed from The Andersen Lab, California Institute of Technology, website[14].

The EEG signals from the dataset were recorded from two healthy male subjects, with 37 and 45 years, using a signal gold plated electrode, placed along the subject's scalp at O1 with linked earlobes reference. The impedances were kept at 5kΩ or lower, and the signal was amplified using a Schwarzer EEG machine set to a sampling rate of 250Hz and filtered using an analog band pass filter at $0.1 - 70$Hz. The signal was then stored to a hard drive using a Brain Data EEG acquisition unit.

The stimuli consisted of two different visual stimuli presented in a pseudo-random order: 75% of the stimuli was the "non-target" stimuli, and consisted of a checkerboard reversing colour; and the other 25% was the "target" stimuli, and consisted of the same checkerboard reversal pattern but with an horizontal and vertical displacement of one-half of the square side length. Subjects were instructed to ignore the non-target stimuli and count the number of appearances of the target stimuli. Signals were recorded for a total of ~2s, corresponding to 256 data points prior and post stimuli, and each session consisted of 200 stimuli presentations. After rejection of trials with artifacts (contamination of the recording with EMG, EOG, etc.) 30 trials from the first subject and 16 trials from the second subject were chosen to be part of the dataset on a total of 46 trials.

Given that our P300 BCI system is built to accept epochs from a more complex stimuli presentation paradigm, we decided to take 600ms of signal prior to stimuli (non-target signal) and 600ms post stimuli (target signal), and use this 1200ms long epoch to evaluate our BCI system.

The P300 BCI system algorithm parameters were adjusted to the dataset, maintaining the intended values of epoch length (i.e. 600ms), the intervals of analysis (i.e. $0 - 300$ms and $300 - 600$ms), and the same feature extraction (i.e. Peak Picking and Area Analysis) and classification methods (i.e. Classes). The only parameters that differed from the ones in our user evaluation were the sampling rate (250Hz vs 128Hz) and the number of classes (2 classes vs 4 classes).

It is worth noting that in this dataset it is not clear how the stimuli were delivered, what type of hardware was used, how long the stimuli were presented, and at what frequency the reversal of the checkerboard occurred.

[14] https://vis.caltech.edu/~rodri/data.htm "Dataset # 4: Pattern visual evoked potentials".

In the following table, we present the accuracy results of our algorithm considering each subject in particular (Table 6.5).

Table 6.5 – Accuracy results from the intra-subject analysis of the P300 dataset, with the developed P300 BCI system algorithm.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	30	26	86.67
2	16	24	87.50
$Mean \pm std$	23.00 ± 9.899	25.00 ± 1.414	87.09 ± 0.5869

From the analysis of Table 6.5, we can see that the two users achieved values of accuracy around 87%, with a minimal value of 86.67% on user #1.

6.1.3. Discussion

The peak average accuracies achieved by the two algorithms in each dataset, considering the rejection of deviant epochs in the SSVEP dataset, are represented in Table 6.6.

Table 6.6 – Highest average accuracies for each dataset.

Dataset	**Peak Average Accuracy (%)**
SSVEP	77.25 ± 5.890
P300	87.09 ± 0.5869

As can be seen in the table above, the P300 BCI system was the algorithm with the highest average accuracy (87.09%), followed by the SSVEP BCI system with a peak average accuracy of 77.25%.

It is worth noting that data in both dataset was collected with distinct paradigms and recording systems, as well as with a different amount of users, with special remark to the P300 dataset that only used two users. So, again, we need to reinforce that these tests were used more as a "proof of concept" than as a mean of evaluating our BCI systems, mostly due to the different recording conditions and stimuli delivery that were used to collect the data. Despite this, we can still see that our algorithm performs well when analysing high quality data collected in an in-lab environment, as is supported by the high values of accuracy achieved in both datasets.

6.2. Evaluation with Users

The performance of our BCI systems was also evaluated in an experimental environment, using the developed UI and volunteer users. In the following sections we present a detailed description of the experimental setup and procedure, as well as the results of this evaluation.

6.2.1. Experimental Setup

All the experiments were performed in a normal office, with adequate lighting and temperature conditions, as well as a minimal comfort for the user, i.e. a desk supporting the LCD screen and a chair at the same level as the screen, in which the user sat. The LCD screen, a 27-inch Dell UltraSharp 2709W widescreen flap panel, had a dimension of $63.21\text{cm} \times 45.19\text{cm} \times 20.00\text{cm}$ (W × H × D), a screen resolution of 1920×1200, and a refresh rate of 60Hz. The chair was positioned at a distance of $\sim 60\text{cm}$ from the screen and all the other people present in the room were asked to make the minimum sound and/or movement near the user as possible, to maximize his/her attention. These less than ideal conditions, in which we recorded the EEG signals, were purposely chosen to mimic a "common daily situation" in which a user may not be in a room with all the lights off [55], or without anyone in the room [55] or even in a shielded room [2]. This type of environment is often called an out-lab environment [58].

6.2.2. Experimental Procedure

The experiments started with the reception to the user, showing where to seat and explaining the general purpose of the study, as well as the way the data was going to be collected (i.e. the Emotiv headset) and that the data would be anonymous. The user was then asked to seat comfortably in the chair, to read and sign an "Informed Consent" form (Annex A) and to start the initial part of a "Self-Assessment Questionnaire" form (Annex B), to collect gender, age, hand preference data and the ingestion of any caffeinated drink (e.g. coffee, tea, energetic drink) in the two hours prior to the study. The appointed researcher explained then how to place the headset, and that the user may experience some discomfort during that study, and that at any moment during the study the user was free to ask to remove the headset and take a break.

The headset was then put on the user's head, and each channel was carefully put in contact with the user's scalp and wet using a saline solution (0.9% NaCl[g/l]) to maintain the electrodes impedance below $5\text{m}\Omega$. The appointed researcher proceeded then by filling the "Experiment Notes" form (Annex C), collecting the type and thickness of the user's hair, the electrode impedance showed by the Emotiv UI and the relative user attention and cooperation throughout the study.

The user interface was then started, and each phase of the interface was explained to the user, prior to the collection of the data:

- T-shaped Maze (Offline): The researcher explained to the user that the avatar had to be moved in the direction of the goal (i.e. yellow coin) and that when the goal is reached, it will be moved to another place and that the user should move the avatar again to the new position, in a succession of eight repetitions. The researcher also made clear that in this phase, the avatar always moves in the right direction, regardless of the intention of the user, which is usually referred to as an Offline environment;

- Full Maze (Online): The researcher explained to the user that, as before, the avatar had to be moved in the direction of the goal and that when the goal is reached, it will be moved to the entrance of the maze again. The researcher also made clear that in this phase, the avatar follows the intention of the user, which is usually referred to as an Online environment, and if there is an error either by the user or by the system, it is the user's responsibility to correct the error.

Between these two phases, the appointed researcher talked to the user to gather the discomfort level and, before changing to a different paradigm (SSVEP and P300), asked the user to fill the correspondent section of the "Self-Assessment Questionnaire" form and explain the differences between each paradigm. Finally, the appointed researcher thanked the user for the cooperation.

Although it was initially planned to test both paradigms in the Offline (T-shaped Maze) and Online (Full Maze) environment, the preliminary tests of the P300 system showed a very low accuracy rate, which translated into the Online environment as an inability to fully control the avatar. Therefore, the T-shaped Maze was common to both the SSVEP and the P300 system, while the Full Maze was only used for the SSVEP system.

6.2.3. Experimental Results

The tests were performed on 15 healthy volunteer subjects. However, data from 5 of the 15 subjects was rejected due to bad electrode connectivity throughout the study (mostly due to the thickness and length of the hair in some of the subjects), continued movement during the recordings (despite being advised not to move during the study), and general lack of concentration and cooperation. It is worth noting that none of the 5 rejected subjects were able to "control" the avatar in the SSVEP online phase of the study, i.e. they did not achieve the goal in the maze and gave up during the experiment.

Thus, the following results, and correspondent analysis, are from the ten remaining subjects, four females and six males age 23.7 ± 3.65 years. We collected data from each subject, recording a single session for each phase, namely: Offline SSVEP; Online SSVEP; and Offline P300. As referred in the last section, data from the Online P300 phase was not collected, due to the constant low accuracy rates during the preliminary tests, which translated into the inability to control the avatar in the Full Maze.

Each of the offline phases, SSVEP and P300, consisted of a 16 decisions session inside a T-shaped maze, while the SSVEP Online phase, consisted of a minimal of a 14 decisions[15] session, without a maximum, i.e. the number of decisions depended on the total number of movements that were done to achieve the intended goal.

[15] The minimal number of decisions needed to achieve the goal, without any errors.

We tested the normality of the data using the D'Agostino & Pearson omnibus normality test, coming to the conclusion that all relevant data was normal ($\alpha = 0.05$). Therefore, all tests considered in the analysis were parametric.

SSVEP

The evaluation of the SSVEP BCI system was split into two phases: an Offline phase, with the purpose of introducing the user to the interface and to its integration with the paradigm, as well as to make an early evaluation of the reaction of the user to the interface, i.e. if the user would be able to successfully control the avatar during the next phase; and an Online phase, with the purpose of evaluating the reaction of the user to a more complex situation, were the decision process needs to incorporate the correction of errors[16]. These evaluations were made using a measure of accuracy per user, defined in the context of this work as the number of correct decisions over the total number of decisions and an overall accuracy, a direct mean of the latter.

In the following tables, we present the accuracy results of the SSVEP Offline phase considering each user in particular (Table 6.7) and all the target frequencies across all users in general (Table 6.8).

Table 6.7 – Accuracy results from the intra-subject analysis of the SSVEP Offline phase.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	16	15	93.75
2	16	14	87.50
3	16	16	100.00
6	16	11	68.75
8	16	15	93.75
9	16	15	93.75
10	16	13	81.25
11	16	14	87.50
14	16	15	93.75
15	16	12	75.00
$Mean \pm std$	16	14.00 ± 1.563	87.50 ± 9.772

Table 6.8 – Accuracy results from the inter-frequency analysis across all users of the SSVEP Offline phase.

Freq. (Hz)	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
10	40	38	95.00
11	40	30	75.00
12	40	39	97.50
13	40	33	82.50
$Mean \pm std$	40	35.00 ± 4.243	87.50 ± 10.61

[16] In the context of this work, we assumed that the users know where they are gazing at, i.e. every error that exists is an error that comes from the BCI system, and not from the user gazing at a point other than that which he needs to look at to achieve the task at hand.

From an analysis of Table 6.7, we can see that users reported a mean value of accuracy of 87.50%, with a minimal value of 68.75% on user #6. It is worth noting that between this phase and the online on, the appointed researcher asked the users if they were able to understand the interface, and users #6 and #15 openly admitted that, in some cases, they gazed at the point correspondent to a decision after the one the interface would require at the time, therefore creating an error that was originated from a bad user decision, and not from the BCI system. In general, these values show that the users demonstrated a minimal perceived comprehension of the mechanisms behind the interface. From an analysis of Table 6.8, we can also see that the accuracies across all used frequencies were similar (around 87.50%), with the lowest accuracies belonging to the 11Hz frequency, corresponding to the "Up" movement, and to the 13Hz frequency, corresponding to the "Down" movement. This information is relevant due to the fact that user #6 reported three perceived errors during a "Down" movement (11Hz) and user #15 reported two perceived errors during an "Up" movement (13Hz).

Passing on to the Online phase, in the following tables, we present the accuracy results of the SSVEP Online phase, again considering each user in particular (Table 6.9) and all the target frequencies across all users in general (Table 6.10).

Table 6.9 – Accuracy results from the intra-subject analysis of the SSVEP Online phase.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	15	14	93.33
2	16	14	87.50
3	14	14	100.00
6	14	14	100.00
8	14	14	100.00
9	25	18	72.00
10	18	15	83.33
11	18	16	88.89
14	18	14	77.78
15	16	14	87.50
$Mean \pm std$	16.80 ± 3.327	14.70 ± 1.337	89.03 ± 9.622

Table 6.10 – Accuracy results from the inter-frequency analysis across all users of the SSVEP Online phase.

Freq. (Hz)	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
10	59	53	89.83
11	23	21	91.30
12	59	53	89.83
13	27	20	74.07
$Mean \pm std$	42.00 ± 19.70	36.75 ± 18.77	86.26 ± 8.154

From an analysis of Table 6.9, we can see that most users took around 17 decisions ($\bar{x}_{\mathrm{TotalEpochs}} = 16.8 \pm 3.32$), with a maximum of 25 for user #9. Since the accuracy of this user is also the lowest (72.00%) we can conclude that most of the excess decisions were errors in which the avatar would not move or, in other words, situations where the BCI system would output a movement that was not feasible (i.e. go through a wall or outside the maze when in the entrance/exit). Another type of error is the one where the BCI system would output an incorrect movement that was possible, but that would deviate from the path which leads to the goal. In this case the user needed to correct this movement, therefore taking an extra two decisions (the first one incorrect, and the second one correct).

Two examples of the difference between these two types of errors are user #1 and user #11. User #1 took 15 decisions to finish the maze, with a total of 14 correct decisions. Since we know that the minimum of decisions needed to finish the maze are 14, we can assume that the extra decision that this user made was an error where no movement was made. Otherwise we would have expected an extra two decisions (one correspondent to the error, and another to correct the error), that can be seen in user #11, that took 18 decisions to finish the maze, with a total of 16 correct decisions, which means that from the extra four decisions, from the 18, two were errors where the BCI system outputted a movement that deviated from the path, and the other two were decisions made to correct this deviation, explained by the extra two correct decisions, from the 16.

In general, users reported a mean value of accuracy of 89.03%, with a minimal value of 72.00% on user #9. From an analysis of Table 6.10, we can also see that the accuracies across all used frequencies were similar (around 89%), with the lowest accuracy belonging to the 13Hz frequency, corresponding to the "Down" movement, most likely due to the different types of errors mentioned above.

SSVEP Offline vs Online

Although there is no significant difference (paired t-test, $\alpha = 0.05$) between the average accuracies of the Offline and Online SSVEP phases (Figure 6.1), most users reported an easier comprehension of the interface and the integration with the paradigm, with some users even managing to improve their accuracy between phases (Figure 6.1**b**). This can be explained by the fact that during the Offline phase, the avatar moved correctly regardless of the user's intent or system errors, and during the Online phase the user would be more aware of the reaction of the interface to their level of attention, allowing for quick correction of gazing mistakes.

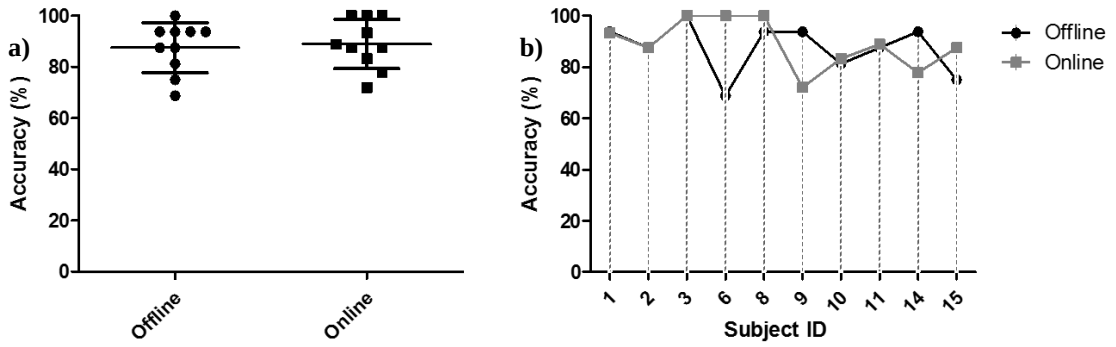


Figure 6.1 – Scatter dot plot **a)** and Grouped dot plot **b)** comparing the accuracy results in the Offline and Online SSVEP phases. **a)** Each circle/square represents a subject accuracy (y-axis), the longer lines represent the sample mean and the two shortest lines represent the standard deviation interval. Subjects are split into the Offline (circle) and Online (square) group (x-axis), which are represented separately. **b)** Each group of circles and squares represent the subject accuracies (y-axis) for the Offline and Online phases, respectively. Data is split per subject (x-axis), with their accuracies being shown for each phase in a grouped manner.

P300

As mentioned, due to issues with the P300 BCI system the evaluation was done using only the Offline phase, again with the purpose of introducing the user to the interface and to its integration with the paradigm, as well as to make an early evaluation of the reaction of the user to the interface, i.e. if the user would be able to successfully control the avatar during the next phase. As before, these evaluations were made using a measure of accuracy per user, defined in the context of this work as the number of correct decisions over the total number of decisions and an overall accuracy, a direct mean of the latter.

In the following table, we present the accuracy results of the P300 Offline phase considering each user in particular (Table 6.11).

Table 6.11 – Accuracy results from the intra-subject analysis of the P300 Offline phase.

Subject ID	**#Total Epochs**	**#Correct Epochs**	**Accuracy (%)**
1	16	3	18.75
2	16	4	25.00
3	16	6	37.50
6	16	2	12.50
8	16	3	18.75
9	16	7	43.75
10	16	5	31.25
11	16	0	0
14	16	8	50.00
15	16	7	43.75
$Mean \pm std$	16	4.50 ± 2.55	28.13 ± 15.93

From an analysis of Table 6.11, we can see that users reported scattered values of accuracy, with a mean of 28.13%, a minimal value of 0% on user #11 and a maximum of 50.00% on user #14. These values are consistent with those observed in the preliminary tests and justify the decision to not evaluate the system in an Online environment. In a more "practical" explanation, the maximum value of 50% is translated to the task at hand as: in each decision point, the P300 BCI system has a 50% chance of outputting the direction that the user was gazing at, i.e. the correct decision, which is too random to allow the user to accurately control the interface and achieve the goal.

Offline SSVEP vs P300

Despite the fact that there is no P300 Online phase, we can still compare the two BCI systems using the Offline phase, that is common to both, and was evaluated under the same conditions. We can observe a significant (paired t-test, $\alpha = 0.05$, p-value< 0.0001) difference between the accuracies of each subject, when comparing the P300 and SSVEP Offline phases, with the latter having a higher accuracy value, as can be seen in Figure 6.2.

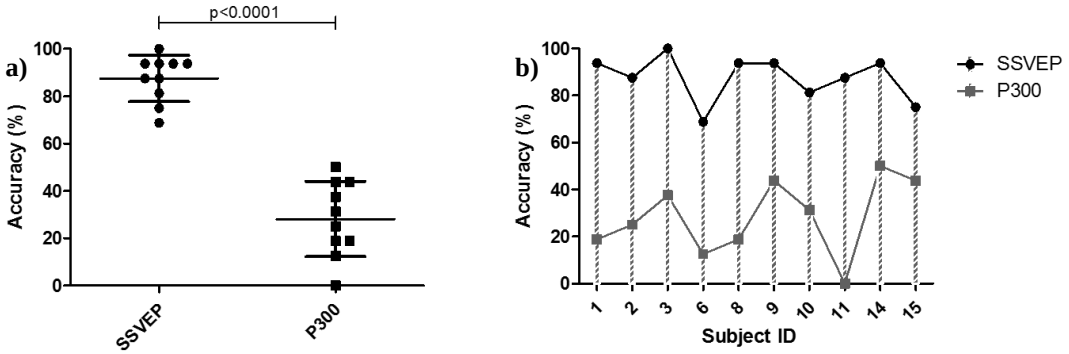


Figure 6.2 – Scatter dot plot **a)** and Grouped dot plot **b)** comparing the accuracy results in the Offline SSVEP and P300 phases. **a)** Each circle/square represents a subject accuracy (y-axis), the longer lines represent the sample mean the two shortest lines represent the standard deviation interval, and the line above the plot represents the p-value result from the paired t-test, $\alpha = 0.05$. Subjects are split into the Offline SSVEP (circle) and Offline P300 (square) group (x-axis), which are represented separately. **b)** Each group of circles and squares represent the subject accuracies (y-axis) for the Offline SSVEP and P300 phases, respectively. Data is split per subject (x-axis), with their accuracies being shown for each phase in a grouped manner.

Self-assessment Questionnaire

In the Self-assessment Questionnaire, users were asked to fill a semantic differential scale for the SSVEP BCI system categorizing it with regards to speed (slow vs fast), concentration (hard vs easy), visual fatigue (tiresome vs relaxing), control level (frustrating vs gratifying) and amusement (boring vs fun), and also for the P300 BCI system, but categorizing it only with regards to visual fatigue and amusement. This discrepancy in the number of categories between both systems is due to the fact that the categories missing in the P300 BCI system are related to the Online phase of the experiment, and therefore were removed from the questionnaire.

The compiled data from the questionnaires is presented in Figure 6.3.

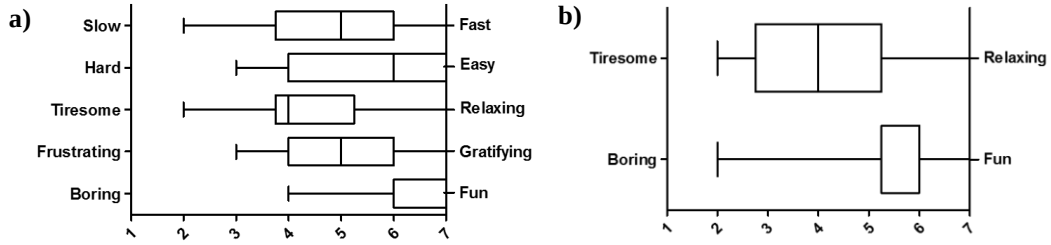


Figure 6.3 – Box and whiskers plots of the Self-assessment questionnaires, regarding the semantic differential scales categorizing the **a)** SSVEP BCI system and the **b)** P300 BCI system. The whiskers represented on the plots are set to minimum and maximum values.

For the SSVEP BCI system (Figure 6.3**a**), most users reported values for concentration (hard vs easy), control level (frustrating vs gratifying) and amusement (boring vs fun) above or equal to the neutral score of 4, except for user #9 that reported a score of 3 for concentration and control level. It is worth noting that user #9 was the user with the lowest accuracy in the Online phase, and the one with the highest number of total decisions made to exit the maze. Most users reported a neutral score of 4 for visual fatigue (tiresome vs relaxing), with some users complaining that the flickering would hurt their eyes after a while. In speed (slow vs fast), users found the system relatively fast (score above neutral of 4) with the exception of two users that found the system slow (score below 3).

Regarding the P300 BCI system (Figure 6.3**b**), most users reported scores of 6 for amusement (boring vs fun), with two users reporting scores below 3 and one user a maximum score of 7, and for visual fatigue (tiresome vs relaxing) users mostly reported scores equal to or below the neutral score of 4, with the exception of three users.

In this questionnaire it was also asked to the users what paradigm they liked the most, to which nine users reported the SSVEP system, and only one reported the P300 system.

It is worth noting that the choices made in this Self-assessment questionnaire are highly influenced due to the fact that the P300 system was not used in an Online scenario, which hinders its "likability" and does not offer an equal opportunity scenario for comparison. Therefore, the analysis of this questionnaire is not very relevant to this work, with regards to the comparison between systems.

SSVEP Online: Concentration vs Accuracy

One of the categories in the Self-assessment Questionnaire was the perceived concentration level of the user, i.e. how "much" did the user had to concentrate to achieve the goal, with a score that ranged from 1–Hard to 7–Easy. Theoretically, this concentration level should be directly related to the accuracy of the SSVEP BCI system, or in other words, the lower the accuracy of the system, the harder the user would need to concentrate to finish the maze, and vice-versa. To assess this relation, we made a correlation analysis and linear regression using the concentration level score and the accuracy of the SSVEP BCI system Online. The results can be seen in Figure 6.4.

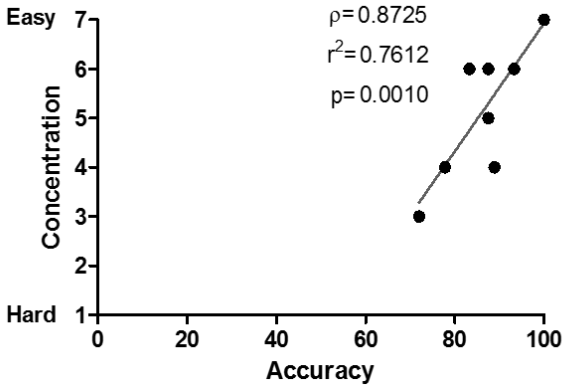


Figure 6.4 – Plot of Concentration level vs Accuracy of the ten users and correspondent correlation and linear regression. ρ – Pearson's correlation coefficient; r^2 – Coefficient of determination or "r squared"; p – p-value.

Correlation quantifies the association between two variables. The correlation coefficient, in this case the parametric Pearson's correlation coefficient or ρ, quantifies the direction and magnitude of the correlation. Our analysis reported a high value of ρ, 0.8725, which means that the two variables tend to increase together, with a strong correlation, and the idea that this correlation is due to random sampling can be discarded due to the small p-value, 0.0010, i.e. the correlation is significant, $\alpha = 0.05$. Furthermore, the r^2 value of 0.7612 means that 76.12% of the observed variance is shared between both variables or in other words, 76.12% of the variance in accuracy can be explained by variation in concentration, and vice-versa.

In summary, the subject concentration level as a direct relation to the accuracy of the SSVEP BCI system, or in other words, the lower the accuracy of the system, the harder the user would need to concentrate to finish the maze, and vice-versa.

6.2.4. Discussion

Accuracies were calculated as the number of correct decisions over the total amount of decisions, per subject, per session. The average accuracies achieved in each phase for all subjects are summarized in Table 6.12.

Table 6.12 – Average accuracy for each phase, considering the ten subjects.

Phase	**Average Accuracy (%)**
Offline SSVEP	87.50 ± 9.772
Online SSVEP	89.03 ± 9.622
Offline P300	28.13 ± 15.93

As can be seen in Table 6.12 , the Online SSVEP phase was the phase with the highest average accuracy (89.03%), followed by the Offline SSVEP phase with a similar accuracy (87.50%) and finally the Offline P300 phase with the lowest accuracy (28.13%).

Since the estimation of the Information Transfer Rate (ITR) only makes sense in a scenario were the user fully controls the BCI, i.e. an Online environment, we can only fully evaluate the SSVEP Online BCI system. Before presenting the result from this evaluation, first we need to describe the inherent features of our BCI system that make this possible, as well as its caveats.

Our BCI falls under the synchronous category, providing the user with information as to when to perform an action, as opposed to asynchronous BCIs which allow users to choose when they wish to control the BCI. The system is also a memory-less and stable discrete transmission channel, with all output commands equally likely to be selected. The classification accuracy is assumed to be the same for all target symbols, as it happens with the equal distribution among all symbols of the classification error. In order to turn this assumption into a more concrete certainty, more tests are required.

The total number of symbols (N) are four, constant throughout the test, the target identification accuracy (P) is given by the accuracy achieved by the system for each subject, and the time needed to output a symbol (T) is $4s$. The value of P is representative of a total of 160 trials, over ten subjects (16 trials per subject). Lastly, our system employed a method for error correction (for one of the two types of possible errors), based on the correction of the erroneous movement of the avatar inside the maze.

Following the methods described in Section 2.3, and using Eq. 2.1, Eq. 2.2 and Eq. 2.3, the ITR achieved by our SSVEP Online BCI system for each subject is presented in Table 6.13, along with an average ITR of 21.01 $bits/min$, representative of our system in the mentioned conditions.

Table 6.13 – ITR results from the intra-subject analysis of the SSVEP Online BCI system. bpm – Bits per minute. † - Due to algorithm constraints, users with an accuracy of 100%, i.e. $P = 1$, were approximated to a P of 0.999.

Subject ID	**Accuracy (%)**	**ITR (bpm)**
1	93.33	23,09
2	87.50	18,87
3	100.00	29,81†
6	100.00	29,81†
8	100.00	29,81†
9	72.00	10,51
10	83.33	16,29
11	88.89	19,81
14	77.78	13,25
15	87.50	18,87
$Mean \pm std$	89.03 ± 9.622	21.01 ± 6.992

This value of ITR is inline with previous works, as we will see in Section 6.3.

6.3. Overall Discussion

Presented in Table 6.14, are the peak average accuracies obtained from our analysis of two (SSVEP and P300) freely available high quality EEG datasets, and from the results of the analysis of data collected from an out-lab environment using data from users, using a commercially affordable portable recording device. It is worth noting that these values are the peak average values from each evaluation, in particular: the SSVEP Datasets value is the one from the evaluation of the Multi-target dataset, considered without the frequency value of 8.2Hz; the SSVEP Users values is from the evaluation of the Online SSVEP phase; the P300 Datasets values is from the evaluation of the P300 dataset; and the P300 Users value is from the evaluation of the Offline SSVEP phase.

Table 6.14 – Peak average accuracy for each paradigm, considering the results from the dataset and user analysis.

Paradigm	**Peak Average Accuracy (%)**	
	Datasets	***Users***
SSVEP	77.25 ± 5.890	89.03 ± 9.622
P300	87.09 ± 0.5869	28.13 ± 15.93

SSVEP

Starting with the SSVEP BCI system, in general our algorithm performed well in the analysis of data collected from users, using a commercially affordable portable recording device, with a peak average accuracy of 89.03% versus a peak average accuracy achieved in the analysis of the high quality SSVEP dataset of 77.25% (Table 6.14). The main differences that can contribute to the lower values in the dataset, and the higher values of accuracy in the data collected from users are:

- **The stimuli delivery system:** Our stimuli delivery system is robust and precise, while there is not much information about the one used by the datasets, which can be a difference to take into account, since the main function of our algorithm relies on frequency separation into classes;
- **The stimuli frequencies:** The used frequencies are tied to the stimuli delivery system, to the refresh rate of the monitor and to the FFT resolution. Our algorithm was developed with our stimuli delivery system in mind, which is independent of the refresh rate of the monitor and uses a very high FFT resolution. This alone can explain the observed issues that our algorithm had with the 8.2Hz frequency from the dataset, and therefore can also be the most important difference to explain the dissimilarity between accuracies;
- **The recording device and recording environment:** In the case of the recording device, there is a common believe among researchers that analog EEG systems perform better than the portable recording devices, although this fact is still contested. On the other hand, EEG signals recorded in an in-lab environment (i.e. dataset) are less prone to artifacts and have a generally higher SNR than those recorded in an out-lab environment (i.e. user's data). Since the observed accuracies were higher in the out-lab environment, using the portable recording device, we believe this differences may not be too important to the observed dissimilarity between accuracies.

On a broader comparison, Table 6.15 summarizes the main works developed on the topic of SSVE-based BCIs (already mentioned in detail on Section 3.1), and our system, offering a summarized comparison view.

Table 6.15 - Summary of the main works developed on the topic of SSVEP-based BCIs. PSDA – Power Spectrum Density Analysis; CCA – Canonical Correlation Analysis; MEC – Minimum Energy Combination; LDA – Linear Discriminant Analysis; N/A – Not Available; ITR – Information Transfer Rate; † - Chumerin et al [58] reportedly rotated the Emotiv headset by 180° so that there were more electrodes in the Occipital area; ǂ - Chumerin et al [58] did not reported an estimate of ITR. All values of accuracy are from an online environment.

	Lalor et al [2]	**Vilic et al [55]**	**Cao et al [57]**	**Chumerin et al [58]**	**Our system**
Electrode System	Wet	Wet	Wet	Dry	**Dry**
Used Electrodes	O1; O2	Oz	POz; P3; P4; Oz; O1; O2	N/A†	**O1; O2**
Portability (device)	No (g.tec)	No (EEG system)	No (EEG system)	Yes (Emotiv)	**Yes (Emotiv)**
Feature Extraction	PSDA	PSDA	CCA	MEC	**PSDA**
Classifier	LDA	Classes	Direct	Direct	**Classes**
Harmonics	No	Yes	N/A	Yes	**Yes**
System	Multiple	Multiple	Multiple	Single	**Single**
Interface	1D "Game"	Speller	Speller	Maze	**Maze**
Accuracy (%)	89.5	90.81	98.78	80.37	**89.03**
ITR (bits/min)	15.5	21.94	61.64	N/Aǂ	**21.01**
#Users	6	9	4	20	**10**

From Table 6.15 we can see that our work is distinctive from other works especially with regards to the type of recording device and interface, and, probably the most outstanding characteristic that is not on the table, with regards to the out-lab recording conditions. The closest work to ours is most likely the one from Chumerin et al [58], that also used the Emotiv recording device, and tested their algorithm with a maze-type interface, similar to ours. But still, they used the Emotiv headset in a peculiar way, rotating it 180°, making it impossible to accurately compare the EEG signals obtained, and also their interface had such an intricate way to be controlled by the user that an ITR measurement would not be feasible. Still we can point out that the accuracy of our system meets the common accuracy from the above mentioned system, despite the different algorithms used. In regards to ITR, this measurement is tied not only to the intrinsic accuracy of the system in question, but also to the way that the interface is built and the stimuli are presented to the user. Taking that into consideration, we can see that our ITR value ($21.01\ bits/min$) is similar to the one from the work of Vilic et al [55] (21.94 bits/min), a work that shares some similarities with ours, with regards to the algorithms used, and slightly larger than the value of the work of Lalor et al [2] (15.5 bits/min), most likely due to the fact that their work was done in a unidimensional interface, hindering the amount of information transferred per minute. Looking at the much higher ITR value from Cao et al [57] (61.64 bits/min), we can see that our system still needs improvements.

P300

Moving to the P300 BCI system, in general our algorithm performed well in the analysis of the high quality P300 dataset, achieving a peak average accuracy of 87.09% (Table 6.14), but performed very poorly in the analysis of data collected from users, using a commercially affordable portable recording device, with a peak average accuracy of 28.13%. The main differences that can contribute to the much lower values in the analysis of data from users, and to the higher values of accuracy in the analysis of the dataset, are:

- **The stimuli delivery system:** Our stimuli delivery system suffered many changes in the course of the development of this work. In spite of that, we believe that the final version is capable of eliciting a P300 response, due to the fact that it was built, taking into account information from various other works in this field. While there is not much information about the stimuli delivery system used by the datasets, our algorithm was able to verify the presence of a P300 response, leading us to take this difference a possible, yet not likely, culprit to the observed bad results;
- **The recording device and recording environment:** In the case of the recording device, we were expecting a commercially affordable portable recording device to perform much worse than the one used in the datasets. This is also true for the out-lab environment versus the in-lab environment used in the dataset. Evoked Potentials are somewhat difficult to obtain and visualize due to normal subject variations, contaminations from outside artifacts, or even just due to the low Signal-to-Noise Ration from an EEG signal. This together with the less than ideal electrode positions used to record the signal, due to constraints inherent to the recording device, the out-lab environment, that most likely contaminated the EEG signal with spurious noise, and the lack of EOG measurement, again contaminating the EEG signal with eye movement artifacts, leads us to believe these to be the most important differences that can explain the huge dissimilarity between accuracies.

On a broader comparison, Table 6.16 summarizes the main works developed on the topic of P300-based BCIs (already mentioned in detail on Section 3.2), and our system, offering a summarized comparison view.

Table 6.16 - Summary of the main works developed on the topic of P300-based BCIs. SWLDA – Step-Wise Linear Discriminant Algorithm; N/A – Not Available; STR. – Single Trial Recognition; VR – Virtual Reality; ITR – Information Transfer Rate; † - In the context of the work of Farwell et al, the ITR is only measured when the user achieves an intended accuracy of 95%; ǂ - Curtin et al reported a bitrate per trial measure, which does not fit into an ITR measure. All values of accuracy are from an online environment, except for our system, which is an offline environment.

		Farwell et al [62]	**Bayliss et al [64]**	**Curtin et al [66]**	**Our system**
Electrode System		Wet	Wet	Wet	**Dry**
Used Electrodes		Pz	Fz; Cz; Pz; P3; P4	FCz; Cz; CP3; CPz; CP4; P3; Pz; P4; Oz	**F3; F4; FC5; FC6**
Portability (device)		No (EEG System)	No (EEG System)	No (EEG System)	**Yes (Emotiv)**
Feature Extraction	**Peak Picking**	Yes	Yes	No	**Yes**
	Area	Yes	No	No	**Yes**
	Correlation	Yes	Yes	No	**No**
	Other	SWLDA	N/A	SWLDA	**N/A**
	Combination	No	No	No	**Yes**
Calibration		Yes	Yes	Yes	**No**
Classifier		Direct	Direct/ Threshold	Direct	**Classes**
Trial		STR	STR	Average	**STR**
System		Multiple	Multiple	Multiple	**Single**
Interface		Speller	VR	Maze	**Maze**
Accuracy (%)		95†	81	88.9	**28.13**
ITR (bits/min)		12.0†	85	N/Aǂ	**N/A**
#Users		4	9	10	**10**

From Table 6.16 we can see that our work is distinctive from every other work. Starting with, probably the most outstanding characteristic that is not on the table, the out-lab recording conditions, that lead us to the eventual choice of a portable recording device, constraining the electrode positions that we had at our disposal, and forcing us to choose less than ideal recording conditions. These conditions also influenced our choices of feature extraction and classification algorithms, leading us to choose algorithms that did not required training time versus the ones that do require it, which are usually more robust algorithms that achieve higher accuracies, but do not fit into the intended "plug-and-play" scenario evaluation. Although, regarding the choice of algorithm, we believe that the P300 BCI system that we developed is robust enough to be able to extract and classify an EEG signal containing P300 signals, as can be verified by the achieved value of accuracy during the evaluation of the P300 datasets, leading us to believe that the problem may lay either in the stimuli delivering system and/or in the recording device.

These differences make it harder to pinpoint the exact culprit in our work, but the bottom line fact remains: our work achieved a much lower accuracy value of 28.13%, in comparison to other works.

6.4. Summary

In this chapter we presented and discussed the two phases of the experimental evaluation of the BCI application, as well as the main results. The purpose of this evaluation was to compare the performance of our BCI systems for a "plug and play" approach using high quality datasets, collected with non-portable EEG recording systems in a lab environment, and using data from users, collected with a commercially affordable portable recording device (i.e. Emotiv) in an out-lab environment.

Regarding the first evaluation, the P300 BCI system was the algorithm with the highest average accuracy (87.09%), followed by the SSVEP BCI system with a peak average accuracy of 77.25%. It is worth noting that these tests were used more as a "proof of concept" than as a mean of evaluating our BCI systems, mostly due to the different recording conditions and stimuli delivery that were used to collect the data. Despite this, we can still see that our algorithm performs well when analysing high quality data collected in an in-lab environment, as is supported by the high values of accuracy achieved in both datasets.

In the second evaluation, the Online SSVEP phase was the phase with the highest average accuracy (89.03%), followed by the Offline SSVEP phase with a similar accuracy (87.50%) and finally the Offline P300 phase with the lowest accuracy (28.13%). The ITR achieved by the Online SSVEP BCI system was 21.01 bits/min.

Finally, we also saw that the subject concentration level has a direct relation to the accuracy of the SSVEP BCI system, or in other words, the lower the accuracy of the system, the harder the user would need to concentrate to finish the maze, and vice-versa.

7. Conclusions and Future Work

The goal of this work was to develop a "plug and play" BCI system, using a commercially affordable portable recording device, to reduce the significant efforts of current BCIs in the fields of setup, calibration and operation. To achieve this goal, we created two BCI systems that allowed for an easy and quick "setup", with no training time ("calibrate") and no user intervention on configurations ("operate"). In order to compare these two systems, we used high quality datasets collected from an in-lab environment using analog EEG systems, and we also developed an interactive interface for spatial navigation, i.e. a maze, in order to evaluate these systems with data collected from voluntary users in an out-lab environment using a portable recording device.

In order to answer these goals, we present the conclusions of this work, discuss his limitations and propose suggestions for future work.

7.1. Conclusions

The main goal of this work was to develop a "plug and play" BCI system, using a commercially affordable portable recording device, to reduce the significant efforts of current BCIs in the fields of setup, calibration and operation. We achieved this goal by creating two distinct BCI systems, based on two exogenous paradigms: an SEP-based system and an ERP-based system. The secondary goal of this work involve the evaluation and comparison of these two distinct systems to identify the best paradigm for a "plug and play" BCI system.

The SEP-based system was developed with a Steady State Visually Evoked Potential (SSVEP) paradigm, without any training time ("calibrate"). The SSVEP-based BCI system was built using a Power Spectral Density Analysis (PSDA) algorithm for the feature extraction system, which considers the fundamental stimuli frequencies and their first two harmonics, and a class based classification system. The EEG signal collected from electrodes $O1$ and $O2$ was used for control

The ERP-based system was developed with a P300 paradigm, which usually requires some calibration setup, but for the purposes of this study, was developed without any training time ("calibrate"). The P300-based BCI system was built using two different algorithms for the feature extraction system, the Peak Picking and Area Analysis algorithms, and the classification system used weighted voting and a class based system to combine the results from these two algorithms. The EEG signal collected from electrodes $F3$, $FC5$, $FC6$ and $F4$ was used for control, and the signal from the remaining electrodes was used to construct a Common Average Referencing (CAR) filter.

In order to fulfil the secondary goal of comparing the two systems and identifying the best for a "plug-and-play" BCI system, we developed a BCI-based interactive interface that is controlled using only inputs from EEG signals recorded with the Emotiv EPOC Headset, which allows for a quick setup time. The interactive interface consists of a 2D maze, in which the BCI system controls the movement in 4 directions of an avatar inside the maze. This control is achieved by having 4 arrows flashing in a distinct way from each other, and for each paradigm. For users to choose the direction they want the avatar to move to, they simply have to gaze at the arrow that represents that direction, and the system will recognize the users' intention, minimizing the operation difficulties.

The SSVEP-based BCI system achieved a peak average accuracy of 89.03%, and a correspondent ITR of 21.01 bits/min, in an Online out-lab environment with data from voluntary users, collected using a commercially affordable portable recording device ("plug-and-play" scenario).

The P300-based BCI system achieved a peak average accuracy of 87.09%, in an Offline in-lab environment with data from datasets, collected using an analog EEG system. Unfortunately, we were unable to fully evaluate the system in an Online out-lab environment with data from voluntary users, collected using a commercially affordable portable recording device ("plug-and-play" scenario), due to the consistent low accuracy values in the Offline run (average of 28.13%).

In conclusion, considering the conditions imposed by this work namely, the out-lab environment, the commercially available recording device, and taking into account the two developed algorithms, the paradigm that obtained better performance in a "plug-and-play" scenario was the SSVEP-based BCI system.

7.2. Limitations and Future Work

Despite the achievement of the mains goals, this work has some limitations. The biggest limitation is tied to the recording device not having the recommended electrode positions for the collection of P300 EEG signals.

In future studies, we firstly suggest the evaluation of the current stimuli/BCI P300 system using signals collected from an analog EEG system. Other possibility is to continue using the stimuli system and the portable recording device and test other feature extraction and classification algorithms, as well as other spatial filtering methods or any other method for signal quality enhancement, for example the simultaneous recording of EOG.

In regards to the stimuli/BCI SSVEP system, we suggest that the SSVEP algorithm in particular needs more robustness and quickness. Perhaps the main course of action in order to achieve this is to perform tests using more stimuli frequencies and more users, to evaluate the response of the algorithm at different frequencies, while also evaluating the robustness of the stimuli system. Another course of action to improve the algorithm is tied to the achieved ITR. A BCI is, in its essence, a peripheral (i.e. a "device that is used to put information into or get information out of the computer") and one of the complaints that users voiced during the experiment was that the interface was slow. This can be improved by shortening the collection times and evaluate the accuracy performance in these situations, maybe even incorporating epoch rejection algorithms, i.e. create a decision threshold, rejecting epochs that have scores below the threshold.

It should be noted, however, that if a "plug-and-play" scenario is still to be taken in mind in future studies (in both systems), the calibration setup should be non-existent or, if really necessary for the considered algorithm, it should be integrated in the interface in such a way that it would not be noted by, or bring any unnecessary configuration setup to the user.

Finally, regarding the developed interactive interface, some users voiced a few complaints about the lack of mazes. The experimental evaluation would be much improved if more mazes are done, possibly even incorporating a random maze generator, with multiple levels of difficulty.

Bibliography

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, Jun. 2002.

[2] E. C. Lalor, S. P. Kelly, C. Finucane, R. Burke, R. Smith, R. B. Reilly, and G. McDarby, "Steady-state VEP-based Brain-computer Interface Control in an Immersive 3D Gaming Environment," *EURASIP J Appl Signal Process*, vol. 2005, pp. 3156–3164, Jan. 2005.

[3] J. N. Mak and J. R. Wolpaw, "Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 187–199, 2009.

[4] N. Birbaumer, C. Weber, C. Neuper, E. Buch, K. Haapen, and L. Cohen, "Physiological regulation of thinking: brain-computer interface (BCI) research," *Prog. Brain Res.*, vol. 159, pp. 369–391, 2006.

[5] "Toward Brain-Computer Interfacing," *MIT Press*. [Online]. Available: https://mitpress.mit.edu/books/toward-brain-computer-interfacing. [Accessed: 13-Sep-2015].

[6] T. M. Vaughan, D. J. McFarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw, "The Wadsworth BCI Research and Development Program: at home with BCI," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 14, no. 2, pp. 229–233, Jun. 2006.

[7] A. Kübler, V. K. Mushahwar, L. R. Hochberg, and J. P. Donoghue, "BCI Meeting 2005--workshop on clinical issues and applications," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 14, no. 2, pp. 131–134, Jun. 2006.

[8] J. R. Wolpaw, G. E. Loeb, B. Z. Allison, E. Donchin, O. F. do Nascimento, W. J. Heetderks, F. Nijboer, W. G. Shain, and J. N. Turner, "BCI meeting 2005-workshop on signals and recording methods," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 138–141, Jun. 2006.

[9] P. R. Kennedy, R. A. Bakay, M. M. Moore, K. Adams, and J. Goldwaithe, "Direct control of a computer from the human central nervous system," *IEEE Trans. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 8, no. 2, pp. 198–202, Jun. 2000.

[10] B. Rebsamen, C. L. Teo, Q. Zeng, V. M. H. Ang, E. Burdet, C. Guan, H. Zhang, and C. Laugier, "Controlling a Wheelchair Indoors Using Thought," *IEEE Intell. Syst.*, vol. 22, no. 2, pp. 18–24, Mar. 2007.

[11] J. J. Vidal, "Toward direct brain-computer communication," *Annu. Rev. Biophys. Bioeng.*, vol. 2, pp. 157–180, 1973.

[12] J. J. Vidal, "Real-time detection of brain events in EEG," *Proc. IEEE*, vol. 65, no. 5, pp. 633–641, May 1977.

[13] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, Jul. 2006.

[14] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An MEG-based brain-computer interface (BCI)," *NeuroImage*, vol. 36, no. 3, pp. 581–593, Jul. 2007.

[15] S. M. Coyle, T. E. Ward, and C. M. Markham, "Brain-computer interface using a simplified functional near-infrared spectroscopy system," *J. Neural Eng.*, vol. 4, no. 3, pp. 219–226, Sep. 2007.

[16] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer, "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 966–970, Jun. 2004.

[17] E. Niedermeyer and F. H. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2005.

[18] A. Searle and L. Kirkup, "A direct comparison of wet, dry and insulating bioelectric recording electrodes," *Physiol. Meas.*, vol. 21, no. 2, pp. 271–283, May 2000.

[19] M. Fernandez and R. Pallas-Areny, “A simple active electrode for power line interference reduction in high resolution biopotential measurements,” in *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine*, 1996, vol. 1, pp. 97–98 vol.1.
[20] F. Z. Padmadinata, J. J. Veerhoek, G. J. A. van Dijk, and J. H. Huijsing, “Microelectronic skin electrode,” *Sens. Actuators B Chem.*, vol. 1, no. 1–6, pp. 491–494, Jan. 1990.
[21] L. A. Geddes, R. Steinberg, and G. Wise, “Dry electrodes and holder for electro-oculography,” *Med. Biol. Eng.*, vol. 11, no. 1, pp. 69–72, Jan. 1973.
[22] A. van Oosterom and J. Strackee, “Computing the lead field of electrodes with axial symmetry,” *Med. Biol. Eng. Comput.*, vol. 21, no. 4, pp. 473–481, Jul. 1983.
[23] B. Alizadeh-Taheri, R. L. Smith, and R. T. Knight, “An Active, Microfabricated, Scalp Electrode-array For EEG Recording,” in *The 8th International Conference on Solid-State Sensors and Actuators, 1995 and Eurosensors IX.. Transducers ’95*, 1995, vol. 1, pp. 67–70.
[24] D. T. Godin, P. A. Parker, and R. N. Scott, “Noise characteristics of stainless-steel surface electrodes,” *Med. Biol. Eng. Comput.*, vol. 29, no. 6, pp. 585–590, Nov. 1991.
[25] J. A. McLaughlin, E. T. McAdams, and J. Anderson, “Novel dry electrode ECG sensor system,” in *Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers*, 1994, p. 804 vol.2-vol.
[26] R. J. Cochran and T. Rosen, “Contact dermatitis caused by ECG electrode paste,” *South. Med. J.*, vol. 73, no. 12, pp. 1667–1668, Dec. 1980.
[27] H. Al-Negheimish, L. Al-Andas, L. Al-Mofeez, A. Al-Abdullatif, N. Al-Khalifa, and A. Al-Wabil, “Brainwave Typing: Comparative Study of P300 and Motor Imagery for Typing Using Dry-Electrode EEG Devices,” in *HCI International 2013 - Posters’ Extended Abstracts*, C. Stephanidis, Ed. Springer Berlin Heidelberg, 2013, pp. 569–573.
[28] Y. Liu, X. Jiang, T. Cao, F. Wan, P. U. Mak, P.-I. Mak, and M. I. Vai, “Implementation of SSVEP based BCI with Emotiv EPOC,” in *2012 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS)*, 2012, pp. 34–37.
[29] W. Mansor, M. S. Abd Rani, and N. Wahy, “Integrating Neural Signal and Embedded System for Controlling Small Motor,” in *Advances in Mechatronics*, H. Martinez-Alfaro, Ed. InTech, 2011.
[30] F. Sharbrough, G. E. Chatrian, R. P. Lesser, H. Luders, M. Nuwer, and T. W. Picton, “American Electroencephalographic Society guidelines for standard electrode position nomenclature,” *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 8, no. 2, pp. 200–202, Apr. 1991.
[31] M. M. Jackson and R. Mappus, “Neural Control Interfaces,” in *Brain-Computer Interfaces*, D. S. Tan and A. Nijholt, Eds. Springer London, 2010, pp. 21–33.
[32] K. E. M. M. PhD and T. F. MD, *Spehlmann’s Evoked Potential Primer, 3e*, 3 edition. Boston: Butterworth-Heinemann, 2001.
[33] E. B. Goldstein, *Encyclopedia of Perception*. SAGE, 2010.
[34] D. Regan, “Some characteristics of average steady-state and transient responses evoked by modulated light,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 20, no. 3, pp. 238–248, Mar. 1966.
[35] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, “Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives,” *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, Apr. 2010.
[36] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu, “Human Cerebral Activation during Steady-State Visual-Evoked Responses,” *J. Neurosci.*, vol. 23, no. 37, pp. 11621–11627, Dec. 2003.
[37] N. R. Galloway, “Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine,” *Br. J. Ophthalmol.*, vol. 74, no. 4, p. 255, Apr. 1990.
[38] D. Purves, *Neuroscience*. Sinauer Associates, 2012.
[39] S. Polyak, *The Vertebrate Visual System*. University of Chicago Press, 1968.
[40] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, “User adaptive BCIs: SSVEP and P300 based interfaces,” *PsychNology J.*, vol. 1, no. 4, pp. 331–354, 2003.
[41] S. Amiri, A. Rabbi, L. Azinfar, and R. Fazel-Rezai, “A Review of P300, SSVEP, and Hybrid P300/SSVEP Brain- Computer Interface Systems,” in *Brain-Computer Interface Systems - Recent Progress and Future Prospects*, R. Fazel-Rezai, Ed. InTech, 2013.

[42] S. J. Luck and E. S. Kappenman, *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, 2011.
[43] J. Polich and C. Margala, "P300 and probability: comparison of oddball and single-stimulus paradigms," *Int. J. Psychophysiol.*, vol. 25, no. 2, pp. 169–176, Feb. 1997.
[44] J. Polich, S. E. Eischen, and G. E. Collins, "P300 from a single auditory stimulus," *Electroencephalogr. Clin. Neurophysiol.*, vol. 92, no. 3, pp. 253–261, May 1994.
[45] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward Enhanced P300 Speller Performance," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 15–21, Jan. 2008.
[46] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999.
[47] S. Bordoloi, U. Sharmah, and S. M. Hazarika, "Motor imagery based BCI for a maze game," in *2012 4th International Conference on Intelligent Human Computer Interaction (IHCI)*, 2012, pp. 1–6.
[48] B. Rockstroh, *Slow cortical potentials and behaviour*. Urban & Schwarzenberg, 1989.
[49] G. J. M. V. Boxtel, K. B. E. Bocker, and G. J. M. VanBoxtel, *Brain and Behavior: Past, Present and Future*, 1 edition. Tilburg: Purdue University Press, 1997.
[50] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: improved accuracy by response verification," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 326–333, Sep. 1998.
[51] P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin, and S. Gao, "A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces," *J. Neural Eng.*, vol. 10, no. 2, p. 26014, Apr. 2013.
[52] Zhonglin Lin, C. Zhang, W. Wu, and X. Gao, "Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.
[53] S. L. Marple, *Digital Spectral Analysis: With Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1986.
[54] W. van Drongelen, *Signal Processing for Neuroscientists: An Introduction to the Analysis of Physiological Signals*. Academic Press, 2006.
[55] A. Vilic, T. W. Kjaer, C. E. Thomsen, S. Puthusserypady, and H. B. D. Sorensen, "DTU BCI speller: an SSVEP-based spelling system with dictionary support," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2013, pp. 2212–2215, 2013.
[56] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, p. 46002, Aug. 2009.
[57] T. Cao, X. Wang, B. Wang, C. M. Wong, F. Wan, P. U. Mak, P. I. Mak, and M. I. Vai, "A high rate online SSVEP based brain-computer interface speller," in *2011 5th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2011, pp. 465–468.
[58] N. Chumerin, N. V. Manyakov, M. van Vliet, A. Robben, A. Combaz, and M. Van Hulle, "Steady-State Visual Evoked Potential-Based Computer Gaming on a Consumer-Grade EEG Device," *IEEE Trans. Comput. Intell. AI Games*, vol. 5, no. 2, pp. 100–110, Jun. 2013.
[59] Y. Wang, Y. t Wang, and T. p Jung, "Visual stimulus design for high-rate SSVEP BCI," *Electron. Lett.*, vol. 46, no. 15, pp. 1057–1058, Jul. 2010.
[60] T. Cao, F. Wan, P. U. Mak, P.-I. Mak, M. I. Vai, and Y. Hu, "Flashing color on the performance of SSVEP-based brain-computer interfaces," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 1819–1822.
[61] J. Cohen and J. Polich, "On the number of trials needed for P300," *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.*, vol. 25, no. 3, pp. 249–255, Apr. 1997.
[62] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.
[63] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 Speller," *J. Neural Eng.*, vol. 3, no. 4, pp. 299–305, Dec. 2006.

[64] J. D. Bayliss, S. A. Inverso, and A. Tentler, "Changing the P300 brain computer interface," *Cyberpsychology Behav. Impact Internet Multimed. Virtual Real. Behav. Soc.*, vol. 7, no. 6, pp. 694–704, Dec. 2004.
[65] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP," *Psychophysiology*, vol. 23, no. 6, pp. 695–703, Nov. 1986.
[66] A. Curtin, H. Ayaz, Y. Liu, P. A. Shewokis, and B. Onaral, "A P300-based EEG-BCI for spatial navigation control," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2012, pp. 3841–3844, 2012.
[67] H. Ayaz, P. A. Shewokis, A. Curtin, M. Izzetoglu, K. Izzetoglu, and B. Onaral, "Using MazeSuite and functional near infrared spectroscopy to study learning in spatial navigation," *J. Vis. Exp. JoVE*, no. 56, 2011.
[68] H. Ayaz, S. L. Allen, S. M. Platek, and B. Onaral, "Maze Suite 1.0: a complete set of tools to prepare, present, and analyze navigational and spatial cognitive neuroscience experiments," *Behav. Res. Methods*, vol. 40, no. 1, pp. 353–359, Feb. 2008.
[69] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.
[70] A. Curtin, H. Ayaz, Y. Liu, P. A. Shewokis, and B. Onaral, "A P300-based EEG-BCI for spatial navigation control," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2012, pp. 3841–3844, 2012.
[71] N. K. Squires, K. C. Squires, and S. A. Hillyard, "Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man," *Electroencephalogr. Clin. Neurophysiol.*, vol. 38, no. 4, pp. 387–401, Apr. 1975.
[72] P. Bocquillon, J.-L. Bourriez, E. Palmero-Soler, A. Destée, L. Defebvre, P. Derambure, and K. Dujardin, "Role of basal ganglia circuits in resisting interference by distracters: a swLORETA study," *PloS One*, vol. 7, no. 3, p. e34239, 2012.
[73] Hiran Ekanayake, "P300 and Emotiv EPOC: Does Emotiv EPOC capture real EEG?," 2010.
[74] R. Q. Quiroga, "Obtaining single stimulus evoked potentials with Wavelet Denoising," *Physica*, vol. 145, pp. 278–292, 2000.
[75] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 115–125, Jan. 2008.

Annex A

BCI-based Virtual Spatial Navigation Control: A Comparison Study

Dissertação para obtenção de Mestrado em Engenharia Biomédica e Biofísica

Formulário de Consentimento Informado

O presente formulário encontra-se integrado no projeto de investigação "BCI-based Virtual Spatial Navigation Control: A Comparison Study" e tem como principal objetivo assegurar que todos os participantes voluntários do estudo estejam devidamente informados sobre os procedimentos e os objetivos do estudo.

Eu (nome completo)_____

1. Tive oportunidade de colocar questões sobre o estudo e esclarecer as minhas dúvidas;
2. Recebi informação que considero suficiente sobre o estudo;
3. Falei com (responsáveis/nome dos investigadores)
   _____.
4. Compreendi que a minha participação neste estudo é voluntária;
5. Compreendi que posso desistir se desejar, sem ter que dar qualquer explicação e sem qualquer repercussão;
6. Declaro não ser portador, nem ter episódios familiares de: epilepsia, convulsões, desmaios de origem desconhecida e/ou transtornos psiquiátricos (e.g.: psicose, esquizofrenia, neurose, etc);
7. Concordo em participar neste estudo de livre vontade.

Assinatura do voluntário	Assinatura do responsável/investigador
_____	_____
Data: ____/____/____	Data: ____/____/____

Annex B

Questionário

ID_____

Questões demográficas:

1. Sexo: Masculino ☐ Feminino ☐
2. Idade: _____
3. É: Destro ☐ Canhoto ☐ Ambi-destro ☐
4. Bebeu alguma bebida cafeinada (café, bebida energética, chá) nas últimas 2h?

   Sim ☐ Não ☐ Se sim, qual e quantas?: _____

(Atenção! Se ainda não fez a 1ª parte da experiência, não prossiga para a próxima parte do questionário. Avise o investigador responsável)

Questões específicas das aplicações:

1. Preencha as seguintes escalas de acordo com a sua opinião, tendo em conta a aplicação (SSVEP) que acabou de testar:

Aborrecida	1	2	3	4	5	6	7	Divertida
Frustante	1	2	3	4	5	6	7	Gratificante
Fatigante	1	2	3	4	5	6	7	Relaxante
Difícil	1	2	3	4	5	6	7	Fácil
Lenta	1	2	3	4	5	6	7	Rápida

(Atenção! Se ainda não fez a 2ª parte da experiência, não prossiga para a próxima parte do questionário. Avise o investigador responsável)

2. Preencha novamente as seguintes escalas de acordo com a sua opinião, mas desta vez tendo em conta a 2ª aplicação (P300) que acabou de testar:

Aborrecida	1	2	3	4	5	6	7	Divertida
Fatigante	1	2	3	4	5	6	7	Relaxante

3. Qual das experiências gostou mais: 1ª ☐ 2ª ☐
4. Aspetos mais positivos / mais negativos:

_____

_____

_____

5. Sugestões/comentários:

_____

_____

_____

Annex C

Notas Experimentais

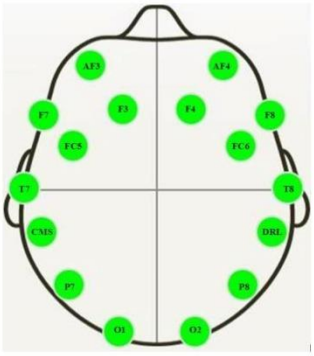
ID_____

1. Cabelo:

Grosso 1 2 3 4 5 6 7 Fino N/A

Comprido 1 2 3 4 5 6 7 Curto N/A

2. Elétrodos:



3. SSVEP:

Desatento/a 1 2 3 4 5 6 7 Focado/a

Indiferente 1 2 3 4 5 6 7 Cooperativo/a

4. P300:

Desatento/a 1 2 3 4 5 6 7 Focado/a

Indiferente 1 2 3 4 5 6 7 Cooperativo/a

5. Notas:

_____

_____

_____